STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact ***the searcher or contact:***

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

- *I am an examiner in Workgroup:* ☐ *Example: 1610*
- *Relevant prior art **found**, search results used as follows:*
  - ☐ 102 rejection
  - ☐ 103 rejection
  - ☐ Cited as being of interest.
  - ☐ Helped examiner better understand the invention.
  - ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
  (journal articles, conference proceedings, new product announcements etc.)

- *Relevant prior art **not found:***
  - ☐ Results verified the lack of relevant prior art (helped determine patentability).
  - ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC Biotech-Chem Library, Remsen Bldg

OM nucleic - nucleic search, using sw model

Run on: April 12, 2006, 13:56:30 ; Search time 803 Seconds
(without alignments)
298.645 Million cell updates/sec

Title: SEQ1-4023-4051-4037A
Perfect score: 29
Sequence: 1 cctctctggagctcaggcatgagccagca 29

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 9793542 seqs, 4134689005 residues

Word size : 15

Total number of hits satisfying chosen parameters: 568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	17	58.6	634	4	US-09-925-065A-577180	Sequence 577180,
c	2	17	58.6	634	4	US-09-925-065A-577181	Sequence 577181,
c	3	17	58.6	634	4	US-09-925-065A-577182	Sequence 577182,
c	4	17	58.6	2808	5	US-10-027-632-112266	Sequence 112266,
c	5	17	58.6	2808	6	US-10-027-632-112266	Sequence 112266,
c	6	16	55.2	201	8	US-10-719-993-24032	Sequence 24032, A
c	7	16	55.2	201	8	US-10-741-600-70028	Sequence 70028, A
	8	16	55.2	299	4	US-09-925-065A-107140	Sequence 107140,
	9	16	55.2	299	4	US-09-925-065A-107142	Sequence 107142,
	10	16	55.2	365	5	US-10-027-632-283657	Sequence 283657,
	11	16	55.2	365	5	US-10-027-632-283658	Sequence 283658,
	12	16	55.2	365	6	US-10-027-632-283657	Sequence 283657,
	13	16	55.2	365	6	US-10-027-632-283658	Sequence 283658,
c	14	16	55.2	370	4	US-09-925-065A-431188	Sequence 431188,
c	15	16	55.2	386	4	US-09-925-065A-159403	Sequence 159403,
c	16	16	55.2	387	4	US-09-925-065A-159399	Sequence 159399,
c	17	16	55.2	391	4	US-09-925-065A-104346	Sequence 104346,
c	18	16	55.2	391	4	US-09-925-065A-159406	Sequence 159406,
c	19	16	55.2	391	8	US-10-357-930-56370	Sequence 56370, A
	20	16	55.2	394	3	US-09-918-995-4596	Sequence 4596, Ap
c	21	16	55.2	395	4	US-09-925-065A-159400	Sequence 159400,
c	22	16	55.2	396	4	US-09-925-065A-173155	Sequence 173155,
c	23	16	55.2	400	4	US-09-925-065A-173157	Sequence 173157,
c	24	16	55.2	400	4	US-09-925-065A-173158	Sequence 173158,
c	25	16	55.2	400	4	US-09-925-065A-173159	Sequence 173159,
c	26	16	55.2	401	4	US-09-925-065A-104348	Sequence 104348,
c	27	16	55.2	401	4	US-09-925-065A-159410	Sequence 159410,
	28	16	55.2	410	8	US-10-357-930-58687	Sequence 58687, A
	29	16	55.2	420	4	US-09-925-065A-107139	Sequence 107139,
	30	16	55.2	420	4	US-09-925-065A-169368	Sequence 169368,
	31	16	55.2	420	4	US-09-925-065A-169369	Sequence 169369,
c	32	16	55.2	423	4	US-09-925-065A-159404	Sequence 159404,
c	33	16	55.2	424	4	US-09-925-065A-104347	Sequence 104347,
c	34	16	55.2	426	4	US-09-925-065A-159401	Sequence 159401,
c	35	16	55.2	426	4	US-09-925-065A-159405	Sequence 159405,
c	36	16	55.2	426	4	US-09-925-065A-159407	Sequence 159407,
c	37	16	55.2	426	4	US-09-925-065A-159408	Sequence 159408,
c	38	16	55.2	426	4	US-09-925-065A-159409	Sequence 159409,
	39	16	55.2	426	8	US-10-357-930-18790	Sequence 18790, A
c	40	16	55.2	429	3	US-09-918-995-1903	Sequence 1903, Ap
c	41	16	55.2	443	3	US-09-918-995-12842	Sequence 12842, A
	42	16	55.2	446	5	US-10-027-632-292293	Sequence 292293,
	43	16	55.2	446	6	US-10-027-632-292293	Sequence 292293,
c	44	16	55.2	450	3	US-09-918-995-11884	Sequence 11884, A
c	45	16	55.2	459	7	US-10-242-535A-58297	Sequence 58297, A
c	46	16	55.2	459	7	US-10-085-783A-58297	Sequence 58297, A
	47	16	55.2	462	4	US-09-925-065A-483021	Sequence 483021,
	48	16	55.2	463	3	US-09-864-761-2863	Sequence 2863, Ap
	49	16	55.2	464	8	US-10-674-124A-23306	Sequence 23306, A
c	50	16	55.2	491	4	US-09-925-065A-30838	Sequence 30838, A
c	51	16	55.2	495	8	US-10-357-930-50418	Sequence 50418, A
c	52	16	55.2	501	4	US-09-925-065A-577939	Sequence 577939,
c	53	16	55.2	520	4	US-09-925-065A-21217	Sequence 21217, A
c	54	16	55.2	527	8	US-10-425-115-59914	Sequence 59914, A
	55	16	55.2	529	4	US-09-925-065A-444539	Sequence 444539,
	56	16	55.2	532	4	US-09-925-065A-745207	Sequence 745207,
	57	16	55.2	532	4	US-09-925-065A-745208	Sequence 745208,
c	58	16	55.2	533	4	US-09-925-065A-476643	Sequence 476643,
c	59	16	55.2	533	4	US-09-925-065A-476644	Sequence 476644,
c	60	16	55.2	533	5	US-10-027-632-161985	Sequence 161985,
c	61	16	55.2	533	5	US-10-027-632-161986	Sequence 161986,
c	62	16	55.2	533	6	US-10-027-632-161985	Sequence 161985,
c	63	16	55.2	533	6	US-10-027-632-161986	Sequence 161986,
	64	16	55.2	541	4	US-09-925-065A-740567	Sequence 740567,
	65	16	55.2	543	5	US-10-027-632-271056	Sequence 271056,
	66	16	55.2	543	6	US-10-027-632-271056	Sequence 271056,
	67	16	55.2	548	4	US-09-925-065A-29550	Sequence 29550, A
	68	16	55.2	549	4	US-09-925-065A-556023	Sequence 556023,
	69	16	55.2	549	4	US-09-925-065A-556024	Sequence 556024,
	70	16	55.2	549	4	US-09-925-065A-841088	Sequence 841088,
	71	16	55.2	551	4	US-09-925-065A-928034	Sequence 928034,
	72	16	55.2	551	4	US-09-925-065A-930207	Sequence 930207,
	73	16	55.2	564	8	US-10-357-930-48600	Sequence 48600, A
	74	16	55.2	565	4	US-09-925-065A-777985	Sequence 777985,
c	75	16	55.2	567	4	US-09-925-065A-556758	Sequence 556758,
c	76	16	55.2	567	4	US-09-925-065A-656344	Sequence 656344,
c	77	16	55.2	568	4	US-09-925-065A-141620	Sequence 141620,
c	78	16	55.2	568	4	US-09-925-065A-141621	Sequence 141621,
c	79	16	55.2	573	9	US-10-779-543-19690	Sequence 19690, A
	80	16	55.2	574	4	US-09-925-065A-799743	Sequence 799743,
c	81	16	55.2	585	4	US-09-925-065A-365382	Sequence 365382,
	82	16	55.2	585	4	US-09-925-065A-892775	Sequence 892775,
	83	16	55.2	591	5	US-10-027-632-265585	Sequence 265585,
	84	16	55.2	591	6	US-10-027-632-265585	Sequence 265585,
c	85	16	55.2	593	4	US-09-925-065A-899570	Sequence 899570,
c	86	16	55.2	594	4	US-09-925-065A-261937	Sequence 261937,
c	87	16	55.2	594	4	US-09-925-065A-261938	Sequence 261938,
	88	16	55.2	594	4	US-09-925-065A-519228	Sequence 519228,
	89	16	55.2	595	4	US-09-925-065A-483020	Sequence 483020,
	90	16	55.2	595	4	US-09-925-065A-483022	Sequence 483022,
	91	16	55.2	598	4	US-09-925-065A-869776	Sequence 869776,
	92	16	55.2	601	9	US-10-893-315-906	Sequence 906, App
	93	16	55.2	606	4	US-09-925-065A-680666	Sequence 680666,
c	94	16	55.2	607	4	US-09-925-065A-737401	Sequence 737401,
	95	16	55.2	610	4	US-09-925-065A-767246	Sequence 767246,
c	96	16	55.2	621	4	US-09-925-065A-765518	Sequence 765518,

c 97	16	55.2	622	4	US-09-925-065A-369908	Sequence 369908,
c 98	16	55.2	622	4	US-09-925-065A-369909	Sequence 369909,
c 99	16	55.2	622	5	US-10-027-632-128018	Sequence 128018,
c 100	16	55.2	622	6	US-10-027-632-128018	Sequence 128018,
101	16	55.2	627	4	US-09-925-065A-560490	Sequence 560490,
c 102	16	55.2	629	4	US-09-925-065A-252792	Sequence 252792,
c 103	16	55.2	629	4	US-09-925-065A-252793	Sequence 252793,
c 104	16	55.2	629	4	US-09-925-065A-252794	Sequence 252794,
c 105	16	55.2	629	4	US-09-925-065A-252795	Sequence 252795,
106	16	55.2	635	4	US-09-925-065A-84844	Sequence 84844, A
107	16	55.2	638	4	US-09-925-065A-800541	Sequence 800541,
c 108	16	55.2	655	4	US-09-925-065A-716468	Sequence 716468,
109	16	55.2	657	4	US-09-925-065A-955783	Sequence 955783,
110	16	55.2	659	4	US-09-925-065A-767245	Sequence 767245,
111	16	55.2	659	4	US-09-925-065A-767247	Sequence 767247,
112	16	55.2	674	4	US-09-925-065A-767248	Sequence 767248,
113	16	55.2	693	4	US-09-925-065A-55079	Sequence 55079, A
114	16	55.2	693	4	US-09-925-065A-55080	Sequence 55080, A
115	16	55.2	745	4	US-09-925-065A-925382	Sequence 925382,
c 116	16	55.2	768	6	US-10-012-697-66	Sequence 66, Appl
c 117	16	55.2	768	9	US-10-779-543-22066	Sequence 22066, A
c 118	16	55.2	875	4	US-09-925-065A-9533	Sequence 9533, Ap
119	16	55.2	1044	4	US-09-925-065A-74190	Sequence 74190, A
c 120	16	55.2	1073	4	US-09-925-065A-697158	Sequence 697158,
c 121	16	55.2	1087	4	US-09-925-065A-678822	Sequence 678822,
c 122	16	55.2	1087	4	US-09-925-065A-678823	Sequence 678823,
c 123	16	55.2	1087	4	US-09-925-065A-678824	Sequence 678824,
c 124	16	55.2	1087	4	US-09-925-065A-678825	Sequence 678825,
125	16	55.2	1148	4	US-09-925-065A-727806	Sequence 727806,
c 126	16	55.2	1201	6	US-10-017-161-2025	Sequence 2025, Ap
c 127	16	55.2	1201	6	US-10-292-798-1671	Sequence 1671, Ap
128	16	55.2	1311	4	US-09-925-065A-704573	Sequence 704573,
129	16	55.2	1373	5	US-10-027-632-207779	Sequence 207779,
130	16	55.2	1373	6	US-10-027-632-207779	Sequence 207779,
c 131	16	55.2	1375	4	US-09-925-065A-79996	Sequence 79996, A
c 132	16	55.2	1375	4	US-09-925-065A-79997	Sequence 79997, A
c 133	16	55.2	1386	6	US-10-369-493-47058	Sequence 47058, A
134	16	55.2	1399	4	US-09-925-065A-717278	Sequence 717278,
c 135	16	55.2	1729	4	US-09-925-065A-714716	Sequence 714716,
136	16	55.2	1989	5	US-10-027-632-97968	Sequence 97968, A
137	16	55.2	1989	5	US-10-027-632-97969	Sequence 97969, A
138	16	55.2	1989	6	US-10-027-632-97968	Sequence 97968, A
139	16	55.2	1989	6	US-10-027-632-97969	Sequence 97969, A
c 140	16	55.2	1989	6	US-10-292-798-1541	Sequence 1541, Ap
c 141	16	55.2	1997	4	US-09-925-065A-29187	Sequence 29187, A
c 142	16	55.2	1997	4	US-09-925-065A-29188	Sequence 29188, A
c 143	16	55.2	2098	4	US-09-925-065A-87472	Sequence 87472, A
144	16	55.2	2147	4	US-09-925-065A-709005	Sequence 709005,
145	16	55.2	2366	5	US-10-027-632-112329	Sequence 112329,
146	16	55.2	2366	6	US-10-027-632-112329	Sequence 112329,
147	16	55.2	2385	5	US-10-027-632-262437	Sequence 262437,
148	16	55.2	2385	5	US-10-027-632-262438	Sequence 262438,
149	16	55.2	2385	5	US-10-027-632-262439	Sequence 262439,
150	16	55.2	2385	5	US-10-027-632-262440	Sequence 262440,
151	16	55.2	2385	6	US-10-027-632-262437	Sequence 262437,
152	16	55.2	2385	6	US-10-027-632-262438	Sequence 262438,
153	16	55.2	2385	6	US-10-027-632-262439	Sequence 262439,
154	16	55.2	2385	6	US-10-027-632-262440	Sequence 262440,
155	16	55.2	2524	4	US-09-925-065A-672433	Sequence 672433,
156	16	55.2	2524	4	US-09-925-065A-718111	Sequence 718111,
157	16	55.2	2743	7	US-10-437-963-45244	Sequence 45244, A
c 158	16	55.2	2995	5	US-10-027-632-112178	Sequence 112178,
c 159	16	55.2	2995	5	US-10-027-632-112179	Sequence 112179,
c 160	16	55.2	2995	6	US-10-027-632-112178	Sequence 112178,
c 161	16	55.2	2995	6	US-10-027-632-112179	Sequence 112179,
162	16	55.2	3187	6	US-10-104-047-1004	Sequence 1004, Ap
c 163	16	55.2	4428	7	US-10-641-643-1109	Sequence 1109, Ap
c 164	16	55.2	12123	3	US-09-764-872-752	Sequence 752, App
165	16	55.2	12394	6	US-10-181-875-10	Sequence 10, Appl
c 166	16	55.2	12542	3	US-09-764-864-1774	Sequence 1774, Ap
c 167	16	55.2	16082	8	US-10-741-600-17975	Sequence 17975, A
168	16	55.2	29829	5	US-10-087-192-694	Sequence 694, App
169	16	55.2	32184	3	US-09-764-891-7300	Sequence 7300, Ap
170	16	55.2	32194	3	US-09-764-891-7028	Sequence 7028, Ap
c 171	16	55.2	32367	5	US-10-158-160A-14	Sequence 14, Appl
172	16	55.2	39522	5	US-10-087-192-538	Sequence 538, App
c 173	16	55.2	41787	7	US-10-322-281-820	Sequence 820, App
174	16	55.2	42571	3	US-09-810-347-3	Sequence 3, Appli
175	16	55.2	42571	5	US-10-224-413-3	Sequence 3, Appli
c 176	16	55.2	45698	3	US-09-984-429-344	Sequence 344, App
177	16	55.2	46215	8	US-10-741-600-17973	Sequence 17973, A
c 178	16	55.2	46215	8	US-10-741-600-17973	Sequence 17973, A
c 179	16	55.2	46625	9	US-10-737-318-47	Sequence 47, Appl
180	16	55.2	53242	5	US-10-087-192-118	Sequence 118, App
181	16	55.2	62124	8	US-10-417-375-82	Sequence 82, Appl
182	16	55.2	66973	5	US-10-087-192-574	Sequence 574, App
183	16	55.2	76670	5	US-10-087-192-2050	Sequence 2050, Ap
c 184	16	55.2	111084	8	US-10-723-860-1627	Sequence 1627, Ap
185	16	55.2	114771	8	US-10-723-860-458	Sequence 458, App
c 186	16	55.2	127508	7	US-10-322-281-158	Sequence 158, App
187	16	55.2	128034	6	US-10-282-174-186	Sequence 186, App
188	16	55.2	128034	6	US-10-282-174-187	Sequence 187, App
189	16	55.2	128034	8	US-10-600-009-186	Sequence 186, App
190	16	55.2	128034	8	US-10-600-009-187	Sequence 187, App
191	16	55.2	130207	5	US-10-087-192-1372	Sequence 1372, Ap
c 192	16	55.2	131673	8	US-10-723-860-4422	Sequence 4422, Ap
c 193	16	55.2	131673	9	US-10-756-149-4416	Sequence 4416, Ap
194	16	55.2	142318	5	US-10-087-192-1504	Sequence 1504, Ap
c 195	16	55.2	144723	5	US-10-087-192-1576	Sequence 1576, Ap
196	16	55.2	152330	5	US-10-087-192-1834	Sequence 1834, Ap
197	16	55.2	160274	9	US-10-893-315-140	Sequence 140, App
198	16	55.2	160300	9	US-10-893-315-151	Sequence 151, App
199	16	55.2	162450	5	US-10-071-179-1	Sequence 1, Appli
200	16	55.2	162450	6	US-10-126-704-1	Sequence 1, Appli
201	16	55.2	162450	10	US-11-132-838-1	Sequence 1, Appli
c 202	16	55.2	166181	8	US-10-723-860-1452	Sequence 1452, Ap
c 203	16	55.2	166181	8	US-10-723-860-3281	Sequence 3281, Ap
204	16	55.2	168174	7	US-10-071-411-63	Sequence 63, Appl
205	16	55.2	168174	8	US-10-914-799-63	Sequence 63, Appl
206	16	55.2	168273	7	US-10-071-411-2	Sequence 2, Appli
207	16	55.2	168273	8	US-10-914-799-2	Sequence 2, Appli
208	16	55.2	179041	7	US-10-741-601-5678	Sequence 5678, Ap
c 209	16	55.2	198522	5	US-10-087-192-244	Sequence 244, App
c 210	16	55.2	202100	6	US-10-282-174-484	Sequence 484, App
c 211	16	55.2	202100	8	US-10-600-009-484	Sequence 484, App
c 212	16	55.2	202814	8	US-10-719-993-6812	Sequence 6812, Ap
213	16	55.2	203127	7	US-10-741-601-5654	Sequence 5654, Ap
214	16	55.2	203127	8	US-10-741-600-17666	Sequence 17666, A
215	16	55.2	203654	3	US-09-820-905-3	Sequence 3, Appli
216	16	55.2	203654	9	US-10-473-338A-3	Sequence 3, Appli
c 217	16	55.2	255439	8	US-10-719-993-6799	Sequence 6799, Ap
c 218	16	55.2	256190	7	US-10-322-281-320	Sequence 320, App
219	16	55.2	260209	6	US-10-265-071-23	Sequence 23, Appl
220	16	55.2	260209	6	US-10-025-966A-23	Sequence 23, Appl
c 221	16	55.2	301692	6	US-10-428-487-11	Sequence 11, Appl
c 222	16	55.2	310268	7	US-10-367-094-195	Sequence 195, App
223	16	55.2	337344	8	US-10-388-838-58	Sequence 58, Appl
224	16	55.2	344548	5	US-10-087-192-334	Sequence 334, App
c 225	16	55.2	380963	9	US-10-737-082-5	Sequence 5, Appli
c 226	16	55.2	380963	9	US-10-765-790-5	Sequence 5, Appli
227	16	55.2	401433	9	US-10-737-082-79	Sequence 79, Appl
228	16	55.2	401433	9	US-10-765-790-79	Sequence 79, Appl
c 229	16	55.2	717651	8	US-10-719-993-6817	Sequence 6817, Ap
230	16	55.2	2731748	7	US-10-297-465A-1	Sequence 1, Appli
c 231	16	55.2	3186778	5	US-10-027-632-174961	Sequence 174961,
c 232	16	55.2	3186778	6	US-10-027-632-174961	Sequence 174961,
c 233	15	51.7	25	10	US-11-060-756-201930	Sequence 201930,
c 234	15	51.7	25	10	US-11-060-756-270294	Sequence 270294,
c 235	15	51.7	201	7	US-10-741-601-24417	Sequence 24417, A
236	15	51.7	210	7	US-10-242-535A-1836	Sequence 1836, Ap
237	15	51.7	210	7	US-10-085-783A-1836	Sequence 1836, Ap
c 238	15	51.7	273	3	US-09-783-590-4174	Sequence 4174, Ap
c 239	15	51.7	301	4	US-09-925-065A-511728	Sequence 511728,
c 240	15	51.7	301	4	US-09-925-065A-511729	Sequence 511729,
241	15	51.7	333	9	US-10-499-065A-231	Sequence 231, App
c 242	15	51.7	336	3	US-09-867-701-7850	Sequence 7850, Ap

c 243	15	51.7	345	4	US-09-925-065A-496756	Sequence 496756,
244	15	51.7	345	7	US-10-389-647-295	Sequence 295, App
c 245	15	51.7	370	9	US-10-779-543-11466	Sequence 11466, A
246	15	51.7	371	4	US-09-925-065A-654890	Sequence 654890,
247	15	51.7	377	4	US-09-925-065A-270241	Sequence 270241,
c 248	15	51.7	387	4	US-09-925-065A-238927	Sequence 238927,
c 249	15	51.7	389	9	US-10-779-543-9651	Sequence 9651, Ap
c 250	15	51.7	395	8	US-10-425-115-148896	Sequence 148896,
251	15	51.7	401	3	US-09-783-590-6415	Sequence 6415, Ap
c 252	15	51.7	402	4	US-09-925-065A-238926	Sequence 238926,
c 253	15	51.7	407	4	US-09-925-065A-950450	Sequence 950450,
254	15	51.7	435	9	US-10-499-065A-232	Sequence 232, App
c 255	15	51.7	447	5	US-10-027-632-77192	Sequence 77192, A
c 256	15	51.7	447	5	US-10-027-632-77193	Sequence 77193, A
c 257	15	51.7	447	5	US-10-027-632-77194	Sequence 77194, A
c 258	15	51.7	447	6	US-10-027-632-77192	Sequence 77192, A
c 259	15	51.7	447	6	US-10-027-632-77193	Sequence 77193, A
c 260	15	51.7	447	6	US-10-027-632-77194	Sequence 77194, A
c 261	15	51.7	470	7	US-10-242-535A-31782	Sequence 31782, A
c 262	15	51.7	470	7	US-10-085-783A-31782	Sequence 31782, A
c 263	15	51.7	471	7	US-10-242-535A-27568	Sequence 27568, A
c 264	15	51.7	471	7	US-10-085-783A-27568	Sequence 27568, A
c 265	15	51.7	481	9	US-10-487-901-1082	Sequence 1082, Ap
c 266	15	51.7	490	5	US-10-027-632-87259	Sequence 87259, A
c 267	15	51.7	490	5	US-10-027-632-87260	Sequence 87260, A
c 268	15	51.7	490	5	US-10-027-632-87261	Sequence 87261, A
c 269	15	51.7	490	5	US-10-027-632-316635	Sequence 316635,
c 270	15	51.7	490	5	US-10-027-632-316636	Sequence 316636,
c 271	15	51.7	490	5	US-10-027-632-316637	Sequence 316637,
c 272	15	51.7	490	6	US-10-027-632-87259	Sequence 87259, A
c 273	15	51.7	490	6	US-10-027-632-87260	Sequence 87260, A
c 274	15	51.7	490	6	US-10-027-632-87261	Sequence 87261, A
c 275	15	51.7	490	6	US-10-027-632-316635	Sequence 316635,
c 276	15	51.7	490	6	US-10-027-632-316636	Sequence 316636,
c 277	15	51.7	490	6	US-10-027-632-316637	Sequence 316637,
c 278	15	51.7	492	8	US-10-357-930-18590	Sequence 18590, A
279	15	51.7	494	4	US-09-925-065A-639136	Sequence 639136,
280	15	51.7	494	4	US-09-925-065A-639137	Sequence 639137,
281	15	51.7	494	4	US-09-925-065A-639138	Sequence 639138,
282	15	51.7	498	4	US-09-925-065A-624913	Sequence 624913,
c 283	15	51.7	499	4	US-09-925-065A-556777	Sequence 556777,
284	15	51.7	501	5	US-10-027-632-291280	Sequence 291280,
285	15	51.7	501	6	US-10-027-632-291280	Sequence 291280,
c 286	15	51.7	506	5	US-10-027-632-12991	Sequence 12991, A
c 287	15	51.7	506	6	US-10-027-632-12991	Sequence 12991, A
288	15	51.7	514	4	US-09-925-065A-831590	Sequence 831590,
289	15	51.7	518	3	US-09-764-877-276	Sequence 276, App
290	15	51.7	518	6	US-10-242-515-276	Sequence 276, App
c 291	15	51.7	521	3	US-09-814-353-1092	Sequence 1092, Ap
c 292	15	51.7	521	3	US-09-814-353-7458	Sequence 7458, Ap
293	15	51.7	525	4	US-09-925-065A-468385	Sequence 468385,
294	15	51.7	525	4	US-09-925-065A-468386	Sequence 468386,
295	15	51.7	526	4	US-09-925-065A-426860	Sequence 426860,
c 296	15	51.7	531	5	US-10-027-632-45725	Sequence 45725, A
c 297	15	51.7	531	6	US-10-027-632-45725	Sequence 45725, A
c 298	15	51.7	535	3	US-09-814-353-13843	Sequence 13843, A
299	15	51.7	535	4	US-09-925-065A-603411	Sequence 603411,
c 300	15	51.7	536	7	US-10-424-599-61811	Sequence 61811, A
c 301	15	51.7	537	4	US-09-925-065A-92727	Sequence 92727, A
c 302	15	51.7	542	4	US-09-925-065A-356119	Sequence 356119,
c 303	15	51.7	542	4	US-09-925-065A-356120	Sequence 356120,
c 304	15	51.7	542	4	US-09-925-065A-356121	Sequence 356121,
305	15	51.7	546	4	US-09-925-065A-50642	Sequence 50642, A
306	15	51.7	547	4	US-09-925-065A-212785	Sequence 212785,
c 307	15	51.7	550	5	US-10-027-632-236031	Sequence 236031,
c 308	15	51.7	550	5	US-10-027-632-236032	Sequence 236032,
c 309	15	51.7	550	6	US-10-027-632-236031	Sequence 236031,
c 310	15	51.7	550	6	US-10-027-632-236032	Sequence 236032,
c 311	15	51.7	553	4	US-09-925-065A-270337	Sequence 270337,
c 312	15	51.7	555	4	US-09-925-065A-481585	Sequence 481585,
c 313	15	51.7	564	4	US-09-925-065A-857894	Sequence 857894,
314	15	51.7	566	4	US-09-925-065A-784618	Sequence 784618,
c 315	15	51.7	567	4	US-09-925-065A-46171	Sequence 46171, A
c 316	15	51.7	569	8	US-10-425-115-156696	Sequence 156696,
c 317	15	51.7	572	8	US-10-357-930-48401	Sequence 48401, A
c 318	15	51.7	575	4	US-09-925-065A-270338	Sequence 270338,
c 319	15	51.7	575	4	US-09-925-065A-270339	Sequence 270339,
320	15	51.7	577	4	US-09-925-065A-386623	Sequence 386623,
c 321	15	51.7	578	5	US-10-027-632-274260	Sequence 274260,
c 322	15	51.7	578	5	US-10-027-632-274261	Sequence 274261,
c 323	15	51.7	578	6	US-10-027-632-274260	Sequence 274260,
c 324	15	51.7	578	6	US-10-027-632-274261	Sequence 274261,
325	15	51.7	580	4	US-09-925-065A-199580	Sequence 199580,
326	15	51.7	581	4	US-09-925-065A-318965	Sequence 318965,
c 327	15	51.7	581	4	US-09-925-065A-319269	Sequence 319269,
328	15	51.7	581	5	US-10-027-632-181654	Sequence 181654,
329	15	51.7	581	5	US-10-027-632-318850	Sequence 318850,
330	15	51.7	581	6	US-10-027-632-181654	Sequence 181654,
331	15	51.7	581	6	US-10-027-632-318850	Sequence 318850,
c 332	15	51.7	582	5	US-10-027-632-95921	Sequence 95921, A
c 333	15	51.7	582	5	US-10-027-632-95922	Sequence 95922, A
c 334	15	51.7	582	6	US-10-027-632-95921	Sequence 95921, A
c 335	15	51.7	582	6	US-10-027-632-95922	Sequence 95922, A
c 336	15	51.7	585	4	US-09-925-065A-347150	Sequence 347150,
c 337	15	51.7	585	4	US-09-925-065A-347151	Sequence 347151,
338	15	51.7	585	4	US-09-925-065A-955975	Sequence 955975,
c 339	15	51.7	589	10	US-11-060-756-3402	Sequence 3402, Ap
c 340	15	51.7	589	10	US-11-060-756-3403	Sequence 3403, Ap
c 341	15	51.7	589	10	US-11-060-756-7674	Sequence 7674, Ap
c 342	15	51.7	589	10	US-11-060-756-7675	Sequence 7675, Ap
343	15	51.7	594	4	US-09-925-065A-680933	Sequence 680933,
344	15	51.7	594	4	US-09-925-065A-799965	Sequence 799965,
345	15	51.7	594	5	US-10-027-632-189845	Sequence 189845,
346	15	51.7	594	6	US-10-027-632-189845	Sequence 189845,
c 347	15	51.7	596	4	US-09-925-065A-674125	Sequence 674125,
c 348	15	51.7	600	4	US-09-925-065A-652274	Sequence 652274,
c 349	15	51.7	600	5	US-10-027-632-140853	Sequence 140853,
c 350	15	51.7	600	5	US-10-027-632-140854	Sequence 140854,
c 351	15	51.7	600	5	US-10-027-632-140855	Sequence 140855,
c 352	15	51.7	600	6	US-10-027-632-140853	Sequence 140853,
c 353	15	51.7	600	6	US-10-027-632-140854	Sequence 140854,
c 354	15	51.7	600	6	US-10-027-632-140855	Sequence 140855,
c 355	15	51.7	600	9	US-10-972-079-55410	Sequence 55410, A
c 356	15	51.7	600	9	US-10-972-079-55411	Sequence 55411, A
357	15	51.7	601	8	US-10-357-930-54577	Sequence 54577, A
c 358	15	51.7	604	4	US-09-925-065A-941993	Sequence 941993,
359	15	51.7	605	4	US-09-925-065A-493602	Sequence 493602,
c 360	15	51.7	606	4	US-09-925-065A-348409	Sequence 348409,
361	15	51.7	608	4	US-09-925-065A-398298	Sequence 398298,
362	15	51.7	608	4	US-09-925-065A-935115	Sequence 935115,
c 363	15	51.7	611	4	US-09-925-065A-887092	Sequence 887092,
c 364	15	51.7	611	4	US-09-925-065A-887093	Sequence 887093,
c 365	15	51.7	614	5	US-10-027-632-92830	Sequence 92830, A
c 366	15	51.7	614	6	US-10-027-632-92830	Sequence 92830, A
367	15	51.7	615	4	US-09-925-065A-21541	Sequence 21541, A
368	15	51.7	617	4	US-09-925-065A-935317	Sequence 935317,
c 369	15	51.7	622	4	US-09-925-065A-298062	Sequence 298062,
370	15	51.7	622	5	US-10-101-464A-281	Sequence 281, App
371	15	51.7	622	9	US-10-864-252-281	Sequence 281, App
372	15	51.7	626	4	US-09-925-065A-920521	Sequence 920521,
c 373	15	51.7	632	9	US-10-487-901-3376	Sequence 3376, Ap
374	15	51.7	635	4	US-09-925-065A-301980	Sequence 301980,
375	15	51.7	635	4	US-09-925-065A-301981	Sequence 301981,
376	15	51.7	635	4	US-09-925-065A-941160	Sequence 941160,
c 377	15	51.7	638	4	US-09-925-065A-793409	Sequence 793409,
c 378	15	51.7	641	5	US-10-027-632-21870	Sequence 21870, A
c 379	15	51.7	641	5	US-10-027-632-21871	Sequence 21871, A
c 380	15	51.7	641	5	US-10-027-632-21872	Sequence 21872, A
c 381	15	51.7	641	6	US-10-027-632-21870	Sequence 21870, A
c 382	15	51.7	641	6	US-10-027-632-21871	Sequence 21871, A
c 383	15	51.7	641	6	US-10-027-632-21872	Sequence 21872, A
c 384	15	51.7	643	4	US-09-925-065A-273160	Sequence 273160,
385	15	51.7	644	4	US-09-925-065A-940112	Sequence 940112,
c 386	15	51.7	651	5	US-10-027-632-128899	Sequence 128899,
c 387	15	51.7	651	6	US-10-027-632-128899	Sequence 128899,
c 388	15	51.7	656	5	US-10-027-632-274262	Sequence 274262,

c 389	15	51.7	656	6	US-10-027-632-274262	Sequence 274262,
c 390	15	51.7	660	4	US-09-925-065A-53162	Sequence 53162, A
c 391	15	51.7	662	4	US-09-925-065A-134359	Sequence 134359,
c 392	15	51.7	667	4	US-09-925-065A-921521	Sequence 921521,
c 393	15	51.7	668	4	US-09-925-065A-934679	Sequence 934679,
c 394	15	51.7	687	5	US-10-027-632-243966	Sequence 243966,
c 395	15	51.7	687	5	US-10-027-632-243967	Sequence 243967,
c 396	15	51.7	687	6	US-10-027-632-243966	Sequence 243966,
c 397	15	51.7	687	6	US-10-027-632-243967	Sequence 243967,
c 398	15	51.7	759	5	US-10-027-632-173172	Sequence 173172,
c 399	15	51.7	759	6	US-10-027-632-173172	Sequence 173172,
c 400	15	51.7	765	4	US-09-925-065A-6229	Sequence 6229, Ap
c 401	15	51.7	783	5	US-10-027-632-146094	Sequence 146094,
c 402	15	51.7	783	6	US-10-027-632-146094	Sequence 146094,
c 403	15	51.7	794	4	US-09-925-065A-701729	Sequence 701729,
c 404	15	51.7	801	3	US-09-814-353-19717	Sequence 19717, A
c 405	15	51.7	805	5	US-10-027-632-157231	Sequence 157231,
c 406	15	51.7	805	5	US-10-027-632-157232	Sequence 157232,
c 407	15	51.7	805	5	US-10-027-632-157233	Sequence 157233,
c 408	15	51.7	805	6	US-10-027-632-157231	Sequence 157231,
c 409	15	51.7	805	6	US-10-027-632-157232	Sequence 157232,
c 410	15	51.7	805	6	US-10-027-632-157233	Sequence 157233,
411	15	51.7	817	5	US-10-027-632-157228	Sequence 157228,
412	15	51.7	817	5	US-10-027-632-157229	Sequence 157229,
413	15	51.7	817	5	US-10-027-632-157230	Sequence 157230,
414	15	51.7	817	6	US-10-027-632-157228	Sequence 157228,
415	15	51.7	817	6	US-10-027-632-157229	Sequence 157229,
416	15	51.7	817	6	US-10-027-632-157230	Sequence 157230,
417	15	51.7	818	5	US-10-027-632-155359	Sequence 155359,
418	15	51.7	818	5	US-10-027-632-155360	Sequence 155360,
419	15	51.7	818	6	US-10-027-632-155359	Sequence 155359,
420	15	51.7	818	6	US-10-027-632-155360	Sequence 155360,
421	15	51.7	896	6	US-10-278-536-238	Sequence 238, App
422	15	51.7	914	7	US-10-412-699B-55	Sequence 55, Appl
c 423	15	51.7	984	5	US-10-027-632-121062	Sequence 121062,
c 424	15	51.7	984	5	US-10-027-632-121063	Sequence 121063,
c 425	15	51.7	984	5	US-10-027-632-121064	Sequence 121064,
c 426	15	51.7	984	6	US-10-027-632-121062	Sequence 121062,
c 427	15	51.7	984	6	US-10-027-632-121063	Sequence 121063,
c 428	15	51.7	984	6	US-10-027-632-121064	Sequence 121064,
c 429	15	51.7	1012	5	US-10-012-600B-108	Sequence 108, App
430	15	51.7	1061	4	US-09-925-065A-554557	Sequence 554557,
431	15	51.7	1061	4	US-09-925-065A-554558	Sequence 554558,
432	15	51.7	1064	5	US-10-027-632-116578	Sequence 116578,
433	15	51.7	1064	5	US-10-027-632-116579	Sequence 116579,
434	15	51.7	1064	6	US-10-027-632-116578	Sequence 116578,
435	15	51.7	1064	6	US-10-027-632-116579	Sequence 116579,
c 436	15	51.7	1132	7	US-10-767-701-10023	Sequence 10023, A
437	15	51.7	1162	5	US-10-027-632-116986	Sequence 116986,
438	15	51.7	1162	5	US-10-027-632-116987	Sequence 116987,
439	15	51.7	1162	6	US-10-027-632-116986	Sequence 116986,
440	15	51.7	1162	6	US-10-027-632-116987	Sequence 116987,
441	15	51.7	1174	5	US-10-027-632-318849	Sequence 318849,
442	15	51.7	1174	6	US-10-027-632-318849	Sequence 318849,
c 443	15	51.7	1472	9	US-10-287-436A-783	Sequence 783, App
444	15	51.7	1503	7	US-10-437-963-68198	Sequence 68198, A
c 445	15	51.7	1545	4	US-09-925-065A-689221	Sequence 689221,
446	15	51.7	1600	3	US-09-753-436-117	Sequence 117, App
447	15	51.7	1600	6	US-10-163-942-117	Sequence 117, App
448	15	51.7	1600	8	US-10-745-115-117	Sequence 117, App
c 449	15	51.7	1616	4	US-09-925-065A-62721	Sequence 62721, A
c 450	15	51.7	1735	3	US-09-974-300-970	Sequence 970, App
451	15	51.7	1916	4	US-09-925-065A-54697	Sequence 54697, A
c 452	15	51.7	1947	4	US-09-925-065A-715529	Sequence 715529,
c 453	15	51.7	1947	4	US-09-925-065A-715530	Sequence 715530,
c 454	15	51.7	2053	9	US-10-450-763-25124	Sequence 25124, A
c 455	15	51.7	2197	7	US-10-424-599-140712	Sequence 140712,
456	15	51.7	2234	9	US-10-450-763-5463	Sequence 5463, Ap
457	15	51.7	2253	4	US-09-925-065A-85304	Sequence 85304, A
458	15	51.7	2253	4	US-09-925-065A-85305	Sequence 85305, A
459	15	51.7	2402	9	US-10-450-763-12253	Sequence 12253, A
c 460	15	51.7	2415	4	US-09-925-065A-76261	Sequence 76261, A
c 461	15	51.7	2460	5	US-10-027-632-103688	Sequence 103688,
c 462	15	51.7	2460	5	US-10-027-632-103689	Sequence 103689,
c 463	15	51.7	2460	6	US-10-027-632-103688	Sequence 103688,
c 464	15	51.7	2460	6	US-10-027-632-103689	Sequence 103689,
c 465	15	51.7	3131	9	US-10-887-553A-509	Sequence 509, App
466	15	51.7	3131	9	US-10-450-763-13304	Sequence 13304, A
c 467	15	51.7	3492	6	US-10-108-260A-1690	Sequence 1690, Ap
c 468	15	51.7	3695	5	US-10-097-340-337	Sequence 337, App
c 469	15	51.7	3695	10	US-11-050-926-337	Sequence 337, App
470	15	51.7	3715	5	US-10-101-464A-887	Sequence 887, App
471	15	51.7	3715	9	US-10-864-252-887	Sequence 887, App
c 472	15	51.7	4211	3	US-09-968-007A-135	Sequence 135, App
c 473	15	51.7	4211	3	US-09-968-007A-447	Sequence 447, App
c 474	15	51.7	4211	3	US-09-968-007A-746	Sequence 746, App
c 475	15	51.7	4211	9	US-10-843-641A-6605	Sequence 6605, Ap
c 476	15	51.7	4211	9	US-10-843-641A-6917	Sequence 6917, Ap
c 477	15	51.7	4211	9	US-10-843-641A-7216	Sequence 7216, Ap
c 478	15	51.7	4280	3	US-09-764-891-8833	Sequence 8833, Ap
c 479	15	51.7	4283	3	US-09-764-891-8834	Sequence 8834, Ap
480	15	51.7	4978	3	US-09-764-887-355	Sequence 355, App
481	15	51.7	4978	5	US-10-073-961-355	Sequence 355, App
482	15	51.7	5757	9	US-10-450-763-10120	Sequence 10120, A
483	15	51.7	11101	10	US-11-023-588-3	Sequence 3, Appli
484	15	51.7	13630	3	US-09-764-868-1369	Sequence 1369, Ap
485	15	51.7	13630	3	US-09-764-875-1220	Sequence 1220, Ap
486	15	51.7	14769	3	US-09-764-891-7092	Sequence 7092, Ap
487	15	51.7	17483	8	US-10-751-606-1	Sequence 1, Appli
c 488	15	51.7	17752	3	US-09-748-127-3	Sequence 3, Appli
c 489	15	51.7	17752	7	US-10-669-693-3	Sequence 3, Appli
c 490	15	51.7	19371	3	US-09-764-891-8394	Sequence 8394, Ap
491	15	51.7	19533	6	US-10-292-798-1713	Sequence 1713, Ap
492	15	51.7	19929	3	US-09-764-891-9967	Sequence 9967, Ap
493	15	51.7	20907	3	US-09-764-891-9966	Sequence 9966, Ap
494	15	51.7	21221	6	US-10-017-161-2067	Sequence 2067, Ap
c 495	15	51.7	22324	7	US-10-322-696-124	Sequence 124, App
c 496	15	51.7	27154	3	US-09-764-891-8396	Sequence 8396, Ap
c 497	15	51.7	29282	8	US-10-719-993-6786	Sequence 6786, Ap
498	15	51.7	31162	8	US-10-719-993-6995	Sequence 6995, Ap
499	15	51.7	32195	3	US-09-764-847-1512	Sequence 1512, Ap
500	15	51.7	32195	5	US-10-092-154-1512	Sequence 1512, Ap

ALIGNMENTS

```
RESULT 1
US-09-925-065A-577180/c
; Sequence 577180, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 577180
;   LENGTH: 634
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-577180
```

```
  Query Match             58.6%;  Score 17;  DB 4;  Length 634;
  Best Local Similarity   100.0%;  Pred. No. 8.1;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db        434 TCAGGCATGAGCCAGCA 418


RESULT 2
US-09-925-065A-577181/c
; Sequence 577181, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 577181
;   LENGTH: 634
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-577181

  Query Match             58.6%;  Score 17;  DB 4;  Length 634;
  Best Local Similarity   100.0%;  Pred. No. 8.1;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db        434 TCAGGCATGAGCCAGCA 418


RESULT 3
US-09-925-065A-577182/c
; Sequence 577182, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 577182
;   LENGTH: 634
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-577182

  Query Match             58.6%;  Score 17;  DB 4;  Length 634;
  Best Local Similarity   100.0%;  Pred. No. 8.1;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db        434 TCAGGCATGAGCCAGCA 418


RESULT 4
US-10-027-632-112266/c
; Sequence 112266, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 112266
;   LENGTH: 2808
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-112266

  Query Match             58.6%;  Score 17;  DB 5;  Length 2808;
  Best Local Similarity   100.0%;  Pred. No. 6.1;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db       2373 TCAGGCATGAGCCAGCA 2357


RESULT 5
US-10-027-632-112266/c
; Sequence 112266, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
```

```
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 112266
;   LENGTH: 2808
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-112266

  Query Match             58.6%;  Score 17;  DB 6;  Length 2808;
  Best Local Similarity   100.0%;  Pred. No. 6.1;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db       2373 TCAGGCATGAGCCAGCA 2357


RESULT 6
US-10-719-993-24032/c
; Sequence 24032, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001496
;  CURRENT APPLICATION NUMBER: US/10/719,993
;  CURRENT FILING DATE:  2003-11-24
;  NUMBER OF SEQ ID NOS: 55342
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24032
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-719-993-24032

  Query Match             55.2%;  Score 16;  DB 8;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db         16 CAGGCATGAGCCAGCA 1


RESULT 7
US-10-741-600-70028/c
; Sequence 70028, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 70028
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-600-70028

  Query Match             55.2%;  Score 16;  DB 8;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        186 CAGGCATGAGCCAGCA 171


RESULT 8
US-09-925-065A-107140
; Sequence 107140, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 107140
;   LENGTH: 299
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-107140

  Query Match             55.2%;  Score 16;  DB 4;  Length 299;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        112 CAGGCATGAGCCAGCA 127


RESULT 9
US-09-925-065A-107142
; Sequence 107142, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 107142
;   LENGTH: 299
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-107142
```

```
  Query Match            55.2%;  Score 16;  DB 4;  Length 299;
  Best Local Similarity  100.0%;  Pred. No. 33;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        112 CAGGCATGAGCCAGCA 127


RESULT 10
US-10-027-632-283657
; Sequence 283657, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 283657
;   LENGTH: 365
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-283657

  Query Match            55.2%;  Score 16;  DB 5;  Length 365;
  Best Local Similarity  100.0%;  Pred. No. 32;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TCTCTGGAGCTCAGGC 18
              ||||||||||||||||
Db         37 TCTCTGGAGCTCAGGC 52


RESULT 11
US-10-027-632-283658
; Sequence 283658, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 283658
;   LENGTH: 365
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-283658

  Query Match            55.2%;  Score 16;  DB 5;  Length 365;
  Best Local Similarity  100.0%;  Pred. No. 32;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TCTCTGGAGCTCAGGC 18
              ||||||||||||||||
Db         37 TCTCTGGAGCTCAGGC 52


RESULT 12
US-10-027-632-283657
; Sequence 283657, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 283657
;   LENGTH: 365
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-283657

  Query Match            55.2%;  Score 16;  DB 6;  Length 365;
  Best Local Similarity  100.0%;  Pred. No. 32;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TCTCTGGAGCTCAGGC 18
              ||||||||||||||||
Db         37 TCTCTGGAGCTCAGGC 52


RESULT 13
US-10-027-632-283658
; Sequence 283658, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
```

```
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 283658
;   LENGTH: 365
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-283658

  Query Match             55.2%;  Score 16;  DB 6;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TCTCTGGAGCTCAGGC 18
              ||||||||||||||||
Db         37 TCTCTGGAGCTCAGGC 52


RESULT 14
US-09-925-065A-431188/c
; Sequence 431188, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 431188
;   LENGTH: 370
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-431188

  Query Match             55.2%;  Score 16;  DB 4;  Length 370;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db         71 CAGGCATGAGCCAGCA 56


RESULT 15
US-09-925-065A-159403/c
; Sequence 159403, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159403
;   LENGTH: 386
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159403

  Query Match             55.2%;  Score 16;  DB 4;  Length 386;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        164 CAGGCATGAGCCAGCA 149


RESULT 16
US-09-925-065A-159399/c
; Sequence 159399, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159399
;   LENGTH: 387
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159399

  Query Match             55.2%;  Score 16;  DB 4;  Length 387;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        165 CAGGCATGAGCCAGCA 150


RESULT 17
```

```
US-09-925-065A-104346/c
; Sequence 104346, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 104346
;   LENGTH: 391
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-104346

  Query Match            55.2%;  Score 16;  DB 4;  Length 391;
  Best Local Similarity  100.0%;  Pred. No. 32;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        169 CAGGCATGAGCCAGCA 154


RESULT 18
US-09-925-065A-159406/c
; Sequence 159406, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159406
;   LENGTH: 391
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159406

  Query Match            55.2%;  Score 16;  DB 4;  Length 391;
  Best Local Similarity  100.0%;  Pred. No. 32;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        169 CAGGCATGAGCCAGCA 154


RESULT 19
US-10-357-930-56370/c
; Sequence 56370, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
;  APPLICANT: Schlegel. Robert
;  APPLICANT:  Endege, Wilson
;  APPLICANT:  Monahan, John
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;  TITLE OF INVENTION:  HUMAN PROSTATE CANCER
;  FILE REFERENCE: MRI-007BCN
;  CURRENT APPLICATION NUMBER: US/10/357,930
;  CURRENT FILING DATE:  2003-02-04
;  PRIOR APPLICATION NUMBER: 09/785,276
;  PRIOR FILING DATE: 2003-02-16
;  PRIOR APPLICATION NUMBER: 60/183,319
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: 60/189,862
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: 60/207,454
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/211,314
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/219,007
;  PRIOR FILING DATE: 2000-07-18
;  PRIOR APPLICATION NUMBER: 60/255,281
;  PRIOR FILING DATE: 2000-12-13
;  NUMBER OF SEQ ID NOS: 62232
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 56370
;   LENGTH: 391
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-357-930-56370

  Query Match            55.2%;  Score 16;  DB 8;  Length 391;
  Best Local Similarity  100.0%;  Pred. No. 32;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        285 CAGGCATGAGCCAGCA 270


RESULT 20
US-09-918-995-4596
; Sequence 4596, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4596
;   LENGTH: 394
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-918-995-4596

  Query Match            55.2%;  Score 16;  DB 3;  Length 394;
  Best Local Similarity  100.0%;  Pred. No. 32;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
```

```
                ||||||||||||||||
Db        175 CAGGCATGAGCCAGCA 190


RESULT 21
US-09-925-065A-159400/c
; Sequence 159400, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159400
;   LENGTH: 395
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159400

  Query Match             55.2%;  Score 16;  DB 4;  Length 395;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        173 CAGGCATGAGCCAGCA 158


RESULT 22
US-09-925-065A-173155/c
; Sequence 173155, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 173155
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-173155

  Query Match             55.2%;  Score 16;  DB 4;  Length 396;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        173 CAGGCATGAGCCAGCA 158


RESULT 23
US-09-925-065A-173157/c
; Sequence 173157, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 173157
;   LENGTH: 400
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-173157

  Query Match             55.2%;  Score 16;  DB 4;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        175 CAGGCATGAGCCAGCA 160


RESULT 24
US-09-925-065A-173158/c
; Sequence 173158, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 173158
;   LENGTH: 400
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-173158
```

Query Match 55.2%; Score 16; DB 4; Length 400;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        175 CAGGCATGAGCCAGCA 160
```

RESULT 25
US-09-925-065A-173159/c
; Sequence 173159, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 173159
; LENGTH: 400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-173159

Query Match 55.2%; Score 16; DB 4; Length 400;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        175 CAGGCATGAGCCAGCA 160
```

RESULT 26
US-09-925-065A-104348/c
; Sequence 104348, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 104348
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-104348

Query Match 55.2%; Score 16; DB 4; Length 401;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        144 CAGGCATGAGCCAGCA 129
```

RESULT 27
US-09-925-065A-159410/c
; Sequence 159410, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159410
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-159410

Query Match 55.2%; Score 16; DB 4; Length 401;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        144 CAGGCATGAGCCAGCA 129
```

RESULT 28
US-10-357-930-58687
; Sequence 58687, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel. Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25

```
;  PRIOR APPLICATION NUMBER: 60/211,314
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/219,007
;  PRIOR FILING DATE: 2000-07-18
;  PRIOR APPLICATION NUMBER: 60/255,281
;  PRIOR FILING DATE: 2000-12-13
;  NUMBER OF SEQ ID NOS: 62232
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 58687
;   LENGTH: 410
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-357-930-58687

  Query Match             55.2%;  Score 16;  DB 8;  Length 410;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        258 CAGGCATGAGCCAGCA 273


RESULT 29
US-09-925-065A-107139
; Sequence 107139, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 107139
;   LENGTH: 420
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-107139

  Query Match             55.2%;  Score 16;  DB 4;  Length 420;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        259 CAGGCATGAGCCAGCA 274


RESULT 30
US-09-925-065A-169368
; Sequence 169368, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 169368
;   LENGTH: 420
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-169368

  Query Match             55.2%;  Score 16;  DB 4;  Length 420;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        259 CAGGCATGAGCCAGCA 274


RESULT 31
US-09-925-065A-169369
; Sequence 169369, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 169369
;   LENGTH: 420
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-169369

  Query Match             55.2%;  Score 16;  DB 4;  Length 420;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        259 CAGGCATGAGCCAGCA 274


RESULT 32
US-09-925-065A-159404/c
; Sequence 159404, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
```

```
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159404
;   LENGTH: 423
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159404

  Query Match             55.2%;  Score 16;  DB 4;  Length 423;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db         169 CAGGCATGAGCCAGCA 154


RESULT 33
US-09-925-065A-104347/c
; Sequence 104347, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 104347
;   LENGTH: 424
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-104347

  Query Match             55.2%;  Score 16;  DB 4;  Length 424;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db         169 CAGGCATGAGCCAGCA 154


RESULT 34
US-09-925-065A-159401/c
; Sequence 159401, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159401
;   LENGTH: 426
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159401

  Query Match             55.2%;  Score 16;  DB 4;  Length 426;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db         169 CAGGCATGAGCCAGCA 154


RESULT 35
US-09-925-065A-159405/c
; Sequence 159405, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159405
;   LENGTH: 426
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159405

  Query Match             55.2%;  Score 16;  DB 4;  Length 426;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db         169 CAGGCATGAGCCAGCA 154


RESULT 36
US-09-925-065A-159407/c
; Sequence 159407, Application US/09925065A
```

```
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159407
;   LENGTH: 426
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159407

  Query Match             55.2%;  Score 16;  DB 4;  Length 426;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        169 CAGGCATGAGCCAGCA 154


RESULT 37
US-09-925-065A-159408/c
; Sequence 159408, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159408
;   LENGTH: 426
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159408

  Query Match             55.2%;  Score 16;  DB 4;  Length 426;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        169 CAGGCATGAGCCAGCA 154


RESULT 38
US-09-925-065A-159409/c
; Sequence 159409, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159409
;   LENGTH: 426
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159409

  Query Match             55.2%;  Score 16;  DB 4;  Length 426;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        169 CAGGCATGAGCCAGCA 154


RESULT 39
US-10-357-930-18790
; Sequence 18790, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
;  APPLICANT: Schlegel. Robert
;  APPLICANT:  Endege, Wilson
;  APPLICANT:  Monahan, John
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;  TITLE OF INVENTION:  HUMAN PROSTATE CANCER
;  FILE REFERENCE: MRI-007BCN
;  CURRENT APPLICATION NUMBER: US/10/357,930
;  CURRENT FILING DATE:  2003-02-04
;  PRIOR APPLICATION NUMBER: 09/785,276
;  PRIOR FILING DATE: 2003-02-16
;  PRIOR APPLICATION NUMBER: 60/183,319
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: 60/189,862
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: 60/207,454
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/211,314
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/219,007
;  PRIOR FILING DATE: 2000-07-18
;  PRIOR APPLICATION NUMBER: 60/255,281
;  PRIOR FILING DATE: 2000-12-13
;  NUMBER OF SEQ ID NOS: 62232
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18790
;   LENGTH: 426
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: 388, 390, 401, 403, 410, 412
;   OTHER INFORMATION: n = A,T,C or G
US-10-357-930-18790

  Query Match             55.2%;  Score 16;  DB 8;  Length 426;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        248 CAGGCATGAGCCAGCA 263


RESULT 40
US-09-918-995-1903/c
; Sequence 1903, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1903
;   LENGTH: 429
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(429)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-1903

  Query Match             55.2%;  Score 16;  DB 3;  Length 429;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        420 CAGGCATGAGCCAGCA 405


RESULT 41
US-09-918-995-12842/c
; Sequence 12842, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12842
;   LENGTH: 443
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(443)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-12842

  Query Match             55.2%;  Score 16;  DB 3;  Length 443;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        413 CAGGCATGAGCCAGCA 398


RESULT 42
US-10-027-632-292293
; Sequence 292293, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 292293
;   LENGTH: 446
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-292293

  Query Match             55.2%;  Score 16;  DB 5;  Length 446;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        311 CAGGCATGAGCCAGCA 326


RESULT 43
US-10-027-632-292293
; Sequence 292293, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
```

```
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 292293
;   LENGTH: 446
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-292293

  Query Match             55.2%;  Score 16;  DB 6;  Length 446;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        311 CAGGCATGAGCCAGCA 326


RESULT 44
US-09-918-995-11884/c
; Sequence 11884, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11884
;   LENGTH: 450
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(450)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-11884

  Query Match             55.2%;  Score 16;  DB 3;  Length 450;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        231 CAGGCATGAGCCAGCA 216


RESULT 45
US-10-242-535A-58297/c
; Sequence 58297, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 58297
;   LENGTH: 459
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5)..(5)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (29)..(30)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-58297

  Query Match             55.2%;  Score 16;  DB 7;  Length 459;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        317 CAGGCATGAGCCAGCA 302


RESULT 46
US-10-085-783A-58297/c
; Sequence 58297, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 58297
;   LENGTH: 459
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5)..(5)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (29)..(30)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-58297

  Query Match             55.2%;  Score 16;  DB 7;  Length 459;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        317 CAGGCATGAGCCAGCA 302


RESULT 47
US-09-925-065A-483021
; Sequence 483021, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
```

```
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 483021
;   LENGTH: 462
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-483021

  Query Match             55.2%;  Score 16;  DB 4;  Length 462;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        106 CAGGCATGAGCCAGCA 121


RESULT 48
US-09-864-761-2863
; Sequence 2863, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2863
;   LENGTH: 463
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL118511.6
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =  14
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 8.9
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL =  12
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL =  10
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL =  11
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL =  11
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL =  34
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL =  16
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL =  10
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =  14
US-09-864-761-2863

  Query Match             55.2%;  Score 16;  DB 3;  Length 463;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAG 23
              ||||||||||||||||
Db         36 GGAGCTCAGGCATGAG 51


RESULT 49
US-10-674-124A-23306
; Sequence 23306, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
;  APPLICANT: INOKO, Hidetoshi
;  APPLICANT:  TAMIYA, Gen
;  TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
;  TITLE OF INVENTION:  GENETIC POLYMORPHISM MARKERS
;  FILE REFERENCE: ORIN-003CIP
;  CURRENT APPLICATION NUMBER: US/10/674,124A
;  CURRENT FILING DATE:  2003-09-26
;  PRIOR APPLICATION NUMBER: 10/257,511
;  PRIOR FILING DATE: 2003-03-07
;  PRIOR APPLICATION NUMBER: PCT/JP00/07621
;  PRIOR FILING DATE: 2000-10-30
;  PRIOR APPLICATION NUMBER: JP2000-112699
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: JP2002-327516
;  PRIOR FILING DATE: 2002-09-28
;  PRIOR APPLICATION NUMBER: JP2002-383869
;  PRIOR FILING DATE: 2002-12-09
;  NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 23306
;   LENGTH: 464
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: AC007768.2_13109
;   FEATURE:
;   OTHER INFORMATION: Located on chromosome 18
;   FEATURE:
;   OTHER INFORMATION: Distance between a terminus base of telomere on
;   OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
;   OTHER INFORMATION: sequence : 24410039
```

```
;   FEATURE:
;   OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
;   OTHER INFORMATION: sequence listing upward to telomere on chrosomal short arm and
;   OTHER INFORMATION: 5'-terminus of this base sequence : 83522
US-10-674-124A-23306

  Query Match             55.2%;  Score 16;  DB 8;  Length 464;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        374 CAGGCATGAGCCAGCA 389


RESULT 50
US-09-925-065A-30838/c
; Sequence 30838, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30838
;   LENGTH: 491
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-30838

  Query Match             55.2%;  Score 16;  DB 4;  Length 491;
  Best Local Similarity   100.0%;  Pred. No. 30;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        296 CAGGCATGAGCCAGCA 281


Search completed: April 12, 2006, 14:10:24
Job time : 821 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: April 12, 2006, 13:47:43 ; Search time 139 Seconds
(without alignments)
370.858 Million cell updates/sec

Title: SEQ1-4023-4051-4037A
Perfect score: 29
Sequence: 1 cctctctggagctcaggcatgagccagca 29

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1303057 seqs, 888780828 residues

Word size : 15

Total number of hits satisfying chosen parameters: 179

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	16	55.2	420	3	US-09-621-976-9373	Sequence 9373, Ap
	2	16	55.2	601	3	US-09-949-016-28977	Sequence 28977, A
c	3	16	55.2	601	3	US-09-949-016-46889	Sequence 46889, A
c	4	16	55.2	601	3	US-09-949-016-46890	Sequence 46890, A
	5	16	55.2	601	3	US-09-949-016-60709	Sequence 60709, A
c	6	16	55.2	601	3	US-09-949-016-78940	Sequence 78940, A
c	7	16	55.2	601	3	US-09-949-016-78941	Sequence 78941, A
c	8	16	55.2	601	3	US-09-949-016-78942	Sequence 78942, A
c	9	16	55.2	601	3	US-09-949-016-78943	Sequence 78943, A
c	10	16	55.2	601	3	US-09-949-016-78944	Sequence 78944, A
c	11	16	55.2	601	3	US-09-949-016-115191	Sequence 115191,
	12	16	55.2	601	3	US-09-949-016-135188	Sequence 135188,
	13	16	55.2	601	3	US-09-949-016-158450	Sequence 158450,
c	14	16	55.2	601	3	US-09-949-016-169176	Sequence 169176,
	15	16	55.2	601	3	US-09-949-016-189478	Sequence 189478,
	16	16	55.2	601	3	US-09-949-016-198187	Sequence 198187,
	17	16	55.2	601	3	US-09-949-016-198188	Sequence 198188,
	18	16	55.2	601	3	US-09-949-016-198189	Sequence 198189,
	19	16	55.2	601	3	US-09-949-016-200290	Sequence 200290,
c	20	16	55.2	601	3	US-09-949-016-201857	Sequence 201857,
c	21	16	55.2	601	3	US-09-949-016-201858	Sequence 201858,
	22	16	55.2	601	3	US-09-949-016-204168	Sequence 204168,
	23	16	55.2	3187	3	US-10-104-047-1004	Sequence 1004, Ap
	24	16	55.2	3810	3	US-09-252-991A-6579	Sequence 6579, Ap
c	25	16	55.2	4428	3	US-09-023-655-1109	Sequence 1109, Ap
	26	16	55.2	5889	2	US-08-463-092B-5	Sequence 5, Appli
	27	16	55.2	5889	2	US-08-462-109A-5	Sequence 5, Appli
	28	16	55.2	5889	2	US-08-460-907B-5	Sequence 5, Appli
	29	16	55.2	5889	3	US-08-463-179A-5	Sequence 5, Appli
	30	16	55.2	5889	3	US-08-461-384B-5	Sequence 5, Appli
c	31	16	55.2	8165	3	US-09-949-016-16816	Sequence 16816, A
	32	16	55.2	12394	3	US-09-488-856A-10	Sequence 10, Appl
	33	16	55.2	14664	3	US-08-836-734E-4	Sequence 4, Appli
c	34	16	55.2	15027	3	US-09-949-016-12660	Sequence 12660, A
c	35	16	55.2	15036	3	US-09-949-016-13351	Sequence 13351, A
	36	16	55.2	15778	3	US-09-949-016-13538	Sequence 13538, A
	37	16	55.2	16438	3	US-09-949-016-16165	Sequence 16165, A
c	38	16	55.2	30324	3	US-09-949-016-16037	Sequence 16037, A
	39	16	55.2	31407	3	US-09-949-016-17359	Sequence 17359, A
	40	16	55.2	33519	3	US-09-949-016-17165	Sequence 17165, A
c	41	16	55.2	37802	3	US-09-949-016-12639	Sequence 12639, A
c	42	16	55.2	38206	3	US-09-949-016-15527	Sequence 15527, A
c	43	16	55.2	41125	3	US-09-949-016-12413	Sequence 12413, A
c	44	16	55.2	41126	3	US-09-949-016-17273	Sequence 17273, A
	45	16	55.2	42571	3	US-09-810-347-3	Sequence 3, Appli
	46	16	55.2	46244	3	US-09-949-016-13508	Sequence 13508, A
c	47	16	55.2	50269	3	US-09-949-016-17598	Sequence 17598, A
	48	16	55.2	51723	3	US-09-949-016-12152	Sequence 12152, A
	49	16	55.2	51723	3	US-09-949-016-16991	Sequence 16991, A
c	50	16	55.2	75295	3	US-09-949-002-575	Sequence 575, App
c	51	16	55.2	75296	3	US-09-949-002-799	Sequence 799, App
	52	16	55.2	81701	3	US-09-949-016-14891	Sequence 14891, A
	53	16	55.2	87863	3	US-09-949-016-14402	Sequence 14402, A
	54	16	55.2	94855	3	US-09-949-016-12264	Sequence 12264, A
c	55	16	55.2	102053	3	US-09-949-016-13025	Sequence 13025, A
c	56	16	55.2	111509	3	US-09-949-016-17379	Sequence 17379, A
c	57	16	55.2	114426	3	US-09-949-016-15078	Sequence 15078, A
c	58	16	55.2	128723	3	US-09-949-016-17533	Sequence 17533, A
c	59	16	55.2	135476	3	US-09-949-016-12611	Sequence 12611, A
c	60	16	55.2	135476	3	US-09-949-016-14413	Sequence 14413, A
c	61	16	55.2	137753	3	US-09-949-016-17404	Sequence 17404, A
c	62	16	55.2	147840	3	US-09-949-016-15236	Sequence 15236, A
	63	16	55.2	162450	3	US-09-345-882-1	Sequence 1, Appli
	64	16	55.2	162450	3	US-10-071-179-1	Sequence 1, Appli
	65	16	55.2	168174	3	US-10-071-411A-63	Sequence 63, Appl
	66	16	55.2	168273	3	US-10-071-411A-2	Sequence 2, Appli
c	67	16	55.2	174639	3	US-09-949-016-16509	Sequence 16509, A
	68	16	55.2	186734	3	US-09-949-016-14870	Sequence 14870, A
	69	16	55.2	193689	3	US-09-949-016-12350	Sequence 12350, A
	70	16	55.2	193689	3	US-09-949-016-13088	Sequence 13088, A
c	71	16	55.2	246444	3	US-09-949-016-13113	Sequence 13113, A
c	72	16	55.2	264665	3	US-09-949-016-13747	Sequence 13747, A
c	73	16	55.2	276237	3	US-09-949-016-17504	Sequence 17504, A
	74	16	55.2	283538	3	US-09-949-016-13506	Sequence 13506, A
c	75	16	55.2	784019	3	US-09-949-016-14033	Sequence 14033, A
c	76	16	55.2	828152	3	US-09-949-016-12777	Sequence 12777, A
	77	15	51.7	414	3	US-09-252-991A-13197	Sequence 13197, A
c	78	15	51.7	520	3	US-09-621-976-14633	Sequence 14633, A
	79	15	51.7	528	3	US-09-252-991A-10320	Sequence 10320, A
	80	15	51.7	601	3	US-09-949-016-19327	Sequence 19327, A
c	81	15	51.7	601	3	US-09-949-016-55793	Sequence 55793, A
	82	15	51.7	601	3	US-09-949-016-62931	Sequence 62931, A
	83	15	51.7	601	3	US-09-949-016-64580	Sequence 64580, A
	84	15	51.7	601	3	US-09-949-016-144995	Sequence 144995,
	85	15	51.7	601	3	US-09-949-016-146821	Sequence 146821,
	86	15	51.7	601	3	US-09-949-016-171836	Sequence 171836,
	87	15	51.7	601	3	US-09-949-016-175797	Sequence 175797,
c	88	15	51.7	601	3	US-09-949-016-183920	Sequence 183920,
c	89	15	51.7	601	3	US-09-949-016-184025	Sequence 184025,
c	90	15	51.7	601	3	US-09-949-016-184130	Sequence 184130,
c	91	15	51.7	601	3	US-09-949-016-184235	Sequence 184235,
c	92	15	51.7	601	3	US-09-949-002-9649	Sequence 9649, Ap
	93	15	51.7	622	3	US-10-101-464A-281	Sequence 281, App
c	94	15	51.7	1563	3	US-09-328-352-2831	Sequence 2831, Ap
	95	15	51.7	1600	2	US-08-487-113D-117	Sequence 117, App
	96	15	51.7	1600	2	US-08-720-420A-117	Sequence 117, App
	97	15	51.7	3715	3	US-10-101-464A-887	Sequence 887, App

98	15	51.7	7026	3	US-09-949-016-15139	Sequence 15139, A
99	15	51.7	11093	3	US-09-949-016-12142	Sequence 12142, A
100	15	51.7	17020	3	US-09-949-016-11818	Sequence 11818, A
101	15	51.7	17021	3	US-09-949-016-13555	Sequence 13555, A
102	15	51.7	17616	3	US-09-949-016-13170	Sequence 13170, A
103	15	51.7	24043	3	US-09-949-016-16104	Sequence 16104, A
c 104	15	51.7	28257	3	US-09-949-016-13076	Sequence 13076, A
c 105	15	51.7	29453	3	US-09-949-016-12939	Sequence 12939, A
c 106	15	51.7	29453	3	US-09-949-016-12940	Sequence 12940, A
c 107	15	51.7	29453	3	US-09-949-016-12941	Sequence 12941, A
c 108	15	51.7	29453	3	US-09-949-016-12942	Sequence 12942, A
109	15	51.7	31111	3	US-09-949-016-15628	Sequence 15628, A
c 110	15	51.7	34531	3	US-09-949-016-14604	Sequence 14604, A
c 111	15	51.7	34855	3	US-09-949-016-13004	Sequence 13004, A
c 112	15	51.7	36093	3	US-09-949-016-14664	Sequence 14664, A
c 113	15	51.7	36093	3	US-09-949-016-14665	Sequence 14665, A
114	15	51.7	37215	3	US-09-949-016-15526	Sequence 15526, A
115	15	51.7	37613	3	US-09-949-016-12669	Sequence 12669, A
116	15	51.7	41965	3	US-09-949-016-13067	Sequence 13067, A
117	15	51.7	42348	3	US-09-949-016-17157	Sequence 17157, A
118	15	51.7	43255	3	US-09-949-016-11909	Sequence 11909, A
119	15	51.7	43795	3	US-08-742-185-101	Sequence 101, App
120	15	51.7	45571	3	US-09-949-016-16262	Sequence 16262, A
c 121	15	51.7	45819	3	US-09-949-002-825	Sequence 825, App
122	15	51.7	46050	3	US-09-820-003C-3	Sequence 3, Appli
123	15	51.7	46718	3	US-09-816-093-3	Sequence 3, Appli
124	15	51.7	47698	3	US-09-949-016-13514	Sequence 13514, A
125	15	51.7	48018	3	US-09-949-016-17573	Sequence 17573, A
126	15	51.7	51403	3	US-09-949-016-15057	Sequence 15057, A
127	15	51.7	53332	3	US-09-801-861-3	Sequence 3, Appli
128	15	51.7	53332	3	US-10-224-562-3	Sequence 3, Appli
129	15	51.7	53332	3	US-10-786-065-3	Sequence 3, Appli
c 130	15	51.7	54000	3	US-09-843-377-11	Sequence 11, Appl
c 131	15	51.7	54033	3	US-09-949-016-12091	Sequence 12091, A
c 132	15	51.7	54033	3	US-09-949-016-14325	Sequence 14325, A
c 133	15	51.7	55851	3	US-09-949-016-13732	Sequence 13732, A
134	15	51.7	56131	3	US-09-949-016-12944	Sequence 12944, A
c 135	15	51.7	56302	3	US-09-949-016-11892	Sequence 11892, A
136	15	51.7	56904	3	US-09-949-016-15501	Sequence 15501, A
137	15	51.7	66428	3	US-09-949-016-12917	Sequence 12917, A
138	15	51.7	68392	3	US-09-949-016-13626	Sequence 13626, A
139	15	51.7	68452	3	US-09-949-016-13305	Sequence 13305, A
140	15	51.7	69763	3	US-09-949-016-12584	Sequence 12584, A
141	15	51.7	94748	3	US-09-949-016-12648	Sequence 12648, A
142	15	51.7	94758	3	US-09-949-016-16741	Sequence 16741, A
c 143	15	51.7	96340	3	US-09-949-016-15863	Sequence 15863, A
c 144	15	51.7	100863	3	US-09-949-016-17031	Sequence 17031, A
145	15	51.7	107458	3	US-09-949-016-15687	Sequence 15687, A
146	15	51.7	110402	3	US-09-949-016-17295	Sequence 17295, A
147	15	51.7	110403	3	US-09-949-016-12741	Sequence 12741, A
c 148	15	51.7	116592	3	US-09-818-512-3	Sequence 3, Appli
c 149	15	51.7	116592	3	US-10-354-065-3	Sequence 3, Appli
c 150	15	51.7	117080	3	US-09-949-016-12627	Sequence 12627, A
c 151	15	51.7	124264	3	US-09-949-016-16396	Sequence 16396, A
152	15	51.7	129658	3	US-09-949-016-17195	Sequence 17195, A
153	15	51.7	136917	3	US-09-949-016-16369	Sequence 16369, A
154	15	51.7	139257	3	US-09-920-671-11	Sequence 11, Appl
155	15	51.7	152331	3	US-09-128-155-16	Sequence 16, Appl
156	15	51.7	177251	3	US-09-949-016-15841	Sequence 15841, A
157	15	51.7	193169	3	US-09-949-016-15091	Sequence 15091, A
c 158	15	51.7	194915	3	US-09-949-016-15584	Sequence 15584, A
159	15	51.7	227390	3	US-09-949-016-12201	Sequence 12201, A
160	15	51.7	227391	3	US-09-949-016-13365	Sequence 13365, A
161	15	51.7	232547	3	US-09-949-016-16603	Sequence 16603, A
162	15	51.7	246230	3	US-09-949-016-17019	Sequence 17019, A
163	15	51.7	246230	3	US-09-949-016-17020	Sequence 17020, A
164	15	51.7	246230	3	US-09-949-016-17021	Sequence 17021, A
165	15	51.7	246230	3	US-09-949-016-17022	Sequence 17022, A
c 166	15	51.7	278866	3	US-09-949-016-13922	Sequence 13922, A
c 167	15	51.7	278866	3	US-09-949-016-13923	Sequence 13923, A
c 168	15	51.7	278866	3	US-09-949-016-13924	Sequence 13924, A
c 169	15	51.7	278866	3	US-09-949-016-13925	Sequence 13925, A
c 170	15	51.7	278866	3	US-09-949-016-13926	Sequence 13926, A
c 171	15	51.7	278866	3	US-09-949-016-14699	Sequence 14699, A
c 172	15	51.7	278866	3	US-09-949-016-14700	Sequence 14700, A
c 173	15	51.7	278866	3	US-09-949-016-14701	Sequence 14701, A
c 174	15	51.7	278866	3	US-09-949-016-14702	Sequence 14702, A
c 175	15	51.7	278866	3	US-09-949-016-14703	Sequence 14703, A
176	15	51.7	325791	3	US-09-768-185A-1	Sequence 1, Appli
c 177	15	51.7	325791	3	US-09-768-185A-1	Sequence 1, Appli
178	15	51.7	373182	3	US-09-949-016-17371	Sequence 17371, A
179	15	51.7	373694	3	US-09-949-016-12062	Sequence 12062, A

ALIGNMENTS

```
RESULT 1
US-09-621-976-9373/c
; Sequence 9373, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 9373
;   LENGTH: 420
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-621-976-9373

  Query Match             55.2%;  Score 16;  DB 3;  Length 420;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        268 CAGGCATGAGCCAGCA 253


RESULT 2
US-09-949-016-28977
; Sequence 28977, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28977
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-28977

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAG 16
```

```
                ||||||||||||||||
Db        410 CCTCTCTGGAGCTCAG 425


RESULT 3
US-09-949-016-46889/c
; Sequence 46889, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46889
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-46889

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db         31 CAGGCATGAGCCAGCA 16


RESULT 4
US-09-949-016-46890/c
; Sequence 46890, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46890
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-46890

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db         30 CAGGCATGAGCCAGCA 15


RESULT 5
US-09-949-016-60709
; Sequence 60709, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 60709
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-60709

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        431 CAGGCATGAGCCAGCA 446


RESULT 6
US-09-949-016-78940/c
; Sequence 78940, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 78940
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-78940

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        565 CAGGCATGAGCCAGCA 550


RESULT 7
US-09-949-016-78941/c
; Sequence 78941, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
```

```
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 78941
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-78941

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        538 CAGGCATGAGCCAGCA 523


RESULT 8
US-09-949-016-78942/c
; Sequence 78942, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 78942
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-78942

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        273 CAGGCATGAGCCAGCA 258


RESULT 9
US-09-949-016-78943/c
; Sequence 78943, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 78943
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-78943

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        188 CAGGCATGAGCCAGCA 173


RESULT 10
US-09-949-016-78944/c
; Sequence 78944, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 78944
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-78944

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        101 CAGGCATGAGCCAGCA 86


RESULT 11
US-09-949-016-115191/c
; Sequence 115191, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
```

```
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 115191
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-115191

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 CTCTGGAGCTCAGGCA 19
              ||||||||||||||||
Db        501 CTCTGGAGCTCAGGCA 486


RESULT 12
US-09-949-016-135188
; Sequence 135188, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 135188
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-135188

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        483 CAGGCATGAGCCAGCA 498


RESULT 13
US-09-949-016-158450
; Sequence 158450, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 158450
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-158450

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAG 16
              ||||||||||||||||
Db        410 CCTCTCTGGAGCTCAG 425


RESULT 14
US-09-949-016-169176/c
; Sequence 169176, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 169176
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-169176

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        460 CAGGCATGAGCCAGCA 445


RESULT 15
US-09-949-016-189478
; Sequence 189478, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 189478
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-189478

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
```

```
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        538 CAGGCATGAGCCAGCA 553


RESULT 16
US-09-949-016-198187
; Sequence 198187, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 198187
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-198187

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        502 CAGGCATGAGCCAGCA 517


RESULT 17
US-09-949-016-198188
; Sequence 198188, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 198188
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-198188

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        432 CAGGCATGAGCCAGCA 447


RESULT 18
US-09-949-016-198189
; Sequence 198189, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 198189
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-198189

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        276 CAGGCATGAGCCAGCA 291


RESULT 19
US-09-949-016-200290
; Sequence 200290, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 200290
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-200290

  Query Match             55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db         62 CAGGCATGAGCCAGCA 77


RESULT 20
US-09-949-016-201857/c
```

```
; Sequence 201857, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 201857
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-201857

  Query Match            55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity  100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        480 CAGGCATGAGCCAGCA 465


RESULT 21
US-09-949-016-201858/c
; Sequence 201858, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 201858
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-201858

  Query Match            55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity  100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        167 CAGGCATGAGCCAGCA 152


RESULT 22
US-09-949-016-204168
; Sequence 204168, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 204168
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-204168

  Query Match            55.2%;  Score 16;  DB 3;  Length 601;
  Best Local Similarity  100.0%;  Pred. No. 14;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db         58 CAGGCATGAGCCAGCA 73


RESULT 23
US-10-104-047-1004
; Sequence 1004, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. 6943241el full length cDNA
;  FILE REFERENCE: H1-A0105
;  CURRENT APPLICATION NUMBER: US/10/104,047
;  CURRENT FILING DATE:  2002-03-25
;  PRIOR APPLICATION NUMBER:
;  PRIOR FILING DATE:
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1004
;   LENGTH: 3187
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-104-047-1004

  Query Match            55.2%;  Score 16;  DB 3;  Length 3187;
  Best Local Similarity  100.0%;  Pred. No. 13;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db       1740 CAGGCATGAGCCAGCA 1755


RESULT 24
US-09-252-991A-6579
; Sequence 6579, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 6579
```

```
;   LENGTH: 3810
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6579

  Query Match             55.2%;  Score 16;  DB 3;  Length 3810;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db         68 CAGGCATGAGCCAGCA 83


RESULT 25
US-09-023-655-1109/c
; Sequence 1109, Application US/09023655
; Patent No. 6607879
;  GENERAL INFORMATION:
;    APPLICANT:  Cocks, Benjamin G.
;    APPLICANT:  Susan G. Stuart
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
;    TITLE OF INVENTION:  EXPRESSION
;    NUMBER OF SEQUENCES:  1508
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/023,655
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0001 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1109:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4428 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g184227
US-09-023-655-1109

  Query Match             55.2%;  Score 16;  DB 3;  Length 4428;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db       3728 CAGGCATGAGCCAGCA 3713


RESULT 26
US-08-463-092B-5
; Sequence 5, Application US/08463092B
; Patent No. 5766880
;  GENERAL INFORMATION:
;    APPLICANT:  Cole, Susan P.C.
;    APPLICANT:  Deeley, Roger G.
;    TITLE OF INVENTION:  ISOLATED NUCLEIC ACID MOLECULES ENCODING
;    TITLE OF INVENTION:  MULTIDRUG RESISTANCE PROTEINS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
;      STREET:  Queen's University at Kingston
;      CITY:  Kingston
;      STATE:  Ontario
;      COUNTRY:  CANADA
;      ZIP:  K7L 3N6
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,092B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/966,923
;      FILING DATE:  27-OCT-1992
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/029,340
;      FILING DATE:  8-MAR-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/141,893
;      FILING DATE:  26-OCT-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/407,207
;      FILING DATE:  20-MAR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steeg, Carol Miernicki
;      REGISTRATION NUMBER:  39,539
;      REFERENCE/DOCKET NUMBER:  Q1546
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (613) 545-2342
;      TELEFAX:  (613) 545-6853
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5889 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  6  4589
US-08-463-092B-5

  Query Match             55.2%;  Score 16;  DB 2;  Length 5889;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db       4644 CAGGCATGAGCCAGCA 4659


RESULT 27
US-08-462-109A-5
; Sequence 5, Application US/08462109A
; Patent No. 5882875
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Cole, Susan P.C.
;    APPLICANT:  Deeley, Roger G.
;    TITLE OF INVENTION:  METHODS FOR IDENTIFYING
;    TITLE OF INVENTION:  MULTIDRUG RESISTANT TUMOR CELLS
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,109A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/966,923
;      FILING DATE:  27-0CT-1992
;      APPLICATION NUMBER:  08/029,340
;      FILING DATE:  8-MAR-1993
;      APPLICATION NUMBER:  08/141,893
;      FILING DATE:  26-0CT-1993
;      APPLICATION NUMBER:  08/407,207
;      FILING DATE:  20-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeConti, Giulio A. Jr.
;      REGISTRATION NUMBER:  31,503
;      REFERENCE/DOCKET NUMBER:  PQI-002CP4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5889 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  6  4589
US-08-462-109A-5

  Query Match             55.2%;  Score 16;  DB 2;  Length 5889;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db       4644 CAGGCATGAGCCAGCA 4659


RESULT 28
US-08-460-907B-5
; Sequence 5, Application US/08460907B
; Patent No. 5891724
;  GENERAL INFORMATION:
;    APPLICANT:  Deeley, Roger G.
;    APPLICANT:  Cole, Susan P.C.
;    TITLE OF INVENTION:  METHODS FOR CONFERRING MULTIDRUG
;    TITLE OF INVENTION:  RESISTANCE ON A CELL
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
;      STREET:  Queen's University at Kingston
;      CITY:  Kingston
;      STATE:  Ontario
;      COUNTRY:  CANADA
;      ZIP:  K7L 3N6
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/460,907B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/966,923
;      FILING DATE:  27-0CT-1992
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/029,340
;      FILING DATE:  8-MAR-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/141,893
;      FILING DATE:  26-0CT-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/407,207
;      FILING DATE:  20-MAR-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steeg, Carol Miernicki
;      REGISTRATION NUMBER:  39,539
;      REFERENCE/DOCKET NUMBER:  Q1551
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (613) 545-2342
;      TELEFAX:  (613) 545-6853
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5889 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  6  4589
US-08-460-907B-5

  Query Match             55.2%;  Score 16;  DB 2;  Length 5889;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db       4644 CAGGCATGAGCCAGCA 4659


RESULT 29
US-08-463-179A-5
; Sequence 5, Application US/08463179A
; Patent No. 6001563
;  GENERAL INFORMATION:
;    APPLICANT:  Cole, Susan P.C.
;    APPLICANT:  Deeley, Roger G.
;    TITLE OF INVENTION:  METHODS FOR IDENTIFYING CHEMOSENSITIZERS
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  ASCII text
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/463,179A
;       FILING DATE:
;       CLASSIFICATION:  536
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/966,923
;       FILING DATE:  27-0CT-1992
;       APPLICATION NUMBER:  08/029,340
;       FILING DATE:  8-MAR-1993
;       APPLICATION NUMBER:  08/141,893
;       FILING DATE:  26-0CT-1993
;       APPLICATION NUMBER:  08/407,207
;       FILING DATE:  20-MAR-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  DeConti, Giulio A. Jr.
;       REGISTRATION NUMBER:  31,503
;       REFERENCE/DOCKET NUMBER:  PQI-002CP8
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 227-7400
;       TELEFAX:  (617) 227-5941
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  5889 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  6  4589
US-08-463-179A-5

  Query Match             55.2%;  Score 16;  DB 3;  Length 5889;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db       4644 CAGGCATGAGCCAGCA 4659


RESULT 30
US-08-461-384B-5
; Sequence 5, Application US/08461384B
; Patent No. 6025473
;  GENERAL INFORMATION:
;     APPLICANT:  Cole, Susan P.C.
;     APPLICANT:  Deeley, Roger G.
;     TITLE OF INVENTION:  MULTIDRUG RESISTANCE PROTEINS
;     NUMBER OF SEQUENCES:  10
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
;       STREET:  Queen's University at Kingston
;       CITY:  Kingston
;       STATE:  Ontario
;       COUNTRY:  CANADA
;       ZIP:  K7L 3N6
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  ASCII text
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/461,384B
;       FILING DATE:  05-JUN-95
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/966,923
;       FILING DATE:  27-0CT-1992
;       APPLICATION NUMBER:  08/029,340
;       FILING DATE:  8-MAR-1993
;       APPLICATION NUMBER:  08/141,893
;       FILING DATE:  26-0CT-1993
;       APPLICATION NUMBER:  08/407,207
;       FILING DATE:  20-MAR-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Steeg, Carol Miernicki
;       REGISTRATION NUMBER:  39,539
;       REFERENCE/DOCKET NUMBER:  Q1547
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (613) 545-2342
;       TELEFAX:  (613) 545-6853
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  5889 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  6  4589
US-08-461-384B-5

  Query Match             55.2%;  Score 16;  DB 3;  Length 5889;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db       4644 CAGGCATGAGCCAGCA 4659


RESULT 31
US-09-949-016-16816/c
; Sequence 16816, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16816
;   LENGTH: 8165
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16816

  Query Match             55.2%;  Score 16;  DB 3;  Length 8165;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db       5539 CAGGCATGAGCCAGCA 5524


RESULT 32
US-09-488-856A-10
; Sequence 10, Application US/09488856A
; Patent No. 6316259
; GENERAL INFORMATION:
;  APPLICANT:      Brett P. Monia
```

```
;  APPLICANT:  Robert McKay
;  APPLICANT:  Madeline M. Butler
;  APPLICANT:  Jacqueline Wyatt
;  TITLE OF INVENTION:      ANTISENSE MODULATION OF GLYCOGEN SYNTHASE KINASE 3 ALPHA EXP
;  FILE REFERENCE:      RTS-0115
;  CURRENT APPLICATION NUMBER: US/09/488,856A
;  CURRENT FILING DATE:  2000-01-21
;  NUMBER OF SEQ ID NOS:      88
; SEQ ID NO 10
;   LENGTH: 12394
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (115)...(397)
;   NAME/KEY: CDS
;   LOCATION: (2438)...(2625)
;   NAME/KEY: CDS
;   LOCATION: (5639)...(5722)
;   NAME/KEY: CDS
;   LOCATION: (5864)...(5974)
;   NAME/KEY: CDS
;   LOCATION: (7902)...(8032)
;   NAME/KEY: CDS
;   LOCATION: (8121)...(8227)
;   NAME/KEY: CDS
;   LOCATION: (9197)...(9294)
;   NAME/KEY: CDS
;   LOCATION: (9375)...(9470)
;   NAME/KEY: CDS
;   LOCATION: (9898)...(10084)
;   NAME/KEY: CDS
;   LOCATION: (10431)...(10523)
;   NAME/KEY: CDS
;   LOCATION: (11713)...(11786)
US-09-488-856A-10

  Query Match             55.2%;  Score 16;  DB 3;  Length 12394;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db       3878 CAGGCATGAGCCAGCA 3893


RESULT 33
US-08-836-734E-4
; Sequence 4, Application US/08836734E
; Patent No. 6846623
; GENERAL INFORMATION:
;  APPLICANT: BECKMANN, JACQUES
;  APPLICANT:  RICHARD, ISABELLE
;  TITLE OF INVENTION: LGMD GENE CODING FOR A CALCIUM DEPENDENT PROTEASE
;  FILE REFERENCE: 960-29 AFMB2628AD/FL/SDU
;  CURRENT APPLICATION NUMBER: US/08/836,734E
;  CURRENT FILING DATE:  1997-07-02
;  PRIOR APPLICATION NUMBER: PCT/EP95/04575
;  PRIOR FILING DATE: 1995-11-21
;  PRIOR APPLICATION NUMBER: EP 94402668.1
;  PRIOR FILING DATE: 1994-11-22
;  NUMBER OF SEQ ID NOS: 116
;  SOFTWARE: MS Word
; SEQ ID NO 4
;   LENGTH: 14664
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(14664)
;   OTHER INFORMATION: /label= Figure 8d
US-08-836-734E-4

  Query Match             55.2%;  Score 16;  DB 3;  Length 14664;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db       7106 CAGGCATGAGCCAGCA 7121


RESULT 34
US-09-949-016-12660/c
; Sequence 12660, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12660
;   LENGTH: 15027
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12660

  Query Match             55.2%;  Score 16;  DB 3;  Length 15027;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        351 CAGGCATGAGCCAGCA 336


RESULT 35
US-09-949-016-13351/c
; Sequence 13351, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13351
;   LENGTH: 15036
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13351

  Query Match             55.2%;  Score 16;  DB 3;  Length 15036;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
```

```
              ||||||||||||||||
Db        351 CAGGCATGAGCCAGCA 336


RESULT 36
US-09-949-016-13538
; Sequence 13538, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13538
;   LENGTH: 15778
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13538

  Query Match             55.2%;  Score 16;  DB 3;  Length 15778;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db      15580 CAGGCATGAGCCAGCA 15595


RESULT 37
US-09-949-016-16165
; Sequence 16165, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16165
;   LENGTH: 16438
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16165

  Query Match             55.2%;  Score 16;  DB 3;  Length 16438;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAG 16
              ||||||||||||||||
Db       5383 CCTCTCTGGAGCTCAG 5398


RESULT 38
US-09-949-016-16037/c
; Sequence 16037, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16037
;   LENGTH: 30324
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16037

  Query Match             55.2%;  Score 16;  DB 3;  Length 30324;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db      25001 CAGGCATGAGCCAGCA 24986


RESULT 39
US-09-949-016-17359
; Sequence 17359, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17359
;   LENGTH: 31407
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(31407)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17359

  Query Match             55.2%;  Score 16;  DB 3;  Length 31407;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db      25032 CAGGCATGAGCCAGCA 25047


RESULT 40
US-09-949-016-17165
```

```
; Sequence 17165, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17165
;   LENGTH: 33519
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17165

  Query Match             55.2%;  Score 16;  DB 3;  Length 33519;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db      11296 CAGGCATGAGCCAGCA 11311


RESULT 41
US-09-949-016-12639/c
; Sequence 12639, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12639
;   LENGTH: 37802
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(37802)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12639

  Query Match             55.2%;  Score 16;  DB 3;  Length 37802;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db      32634 CAGGCATGAGCCAGCA 32619


RESULT 42
US-09-949-016-15527/c
; Sequence 15527, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15527
;   LENGTH: 38206
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-15527

  Query Match             55.2%;  Score 16;  DB 3;  Length 38206;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db      33581 CAGGCATGAGCCAGCA 33566


RESULT 43
US-09-949-016-12413/c
; Sequence 12413, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12413
;   LENGTH: 41125
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12413

  Query Match             55.2%;  Score 16;  DB 3;  Length 41125;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db      30521 CAGGCATGAGCCAGCA 30506


RESULT 44
US-09-949-016-17273/c
; Sequence 17273, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

```
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17273
;   LENGTH: 41126
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17273

  Query Match             55.2%;  Score 16;  DB 3;  Length 41126;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db      30521 CAGGCATGAGCCAGCA 30506


RESULT 45
US-09-810-347-3
; Sequence 3, Application US/09810347
; Patent No. 6461847
; GENERAL INFORMATION:
;  APPLICANT: YE, Jane et al.
;  TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001169
;  CURRENT APPLICATION NUMBER: US/09/810,347
;  CURRENT FILING DATE:  2001-03-19
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 42571
;   TYPE: DNA
;   ORGANISM: Human
US-09-810-347-3

  Query Match             55.2%;  Score 16;  DB 3;  Length 42571;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db      32458 CAGGCATGAGCCAGCA 32473


RESULT 46
US-09-949-016-13508
; Sequence 13508, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13508
;   LENGTH: 46244
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13508

  Query Match             55.2%;  Score 16;  DB 3;  Length 46244;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db      11047 CAGGCATGAGCCAGCA 11062


RESULT 47
US-09-949-016-17598/c
; Sequence 17598, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17598
;   LENGTH: 50269
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17598

  Query Match             55.2%;  Score 16;  DB 3;  Length 50269;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db      20413 CAGGCATGAGCCAGCA 20398


RESULT 48
US-09-949-016-12152
; Sequence 12152, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12152
;   LENGTH: 51723
;   TYPE: DNA
;   ORGANISM: Human
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(51723)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12152

  Query Match            55.2%;  Score 16;  DB 3;  Length 51723;
  Best Local Similarity  100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db       8117 CAGGCATGAGCCAGCA 8132


RESULT 49
US-09-949-016-16991
; Sequence 16991, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16991
;   LENGTH: 51723
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(51723)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16991

  Query Match            55.2%;  Score 16;  DB 3;  Length 51723;
  Best Local Similarity  100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db       8117 CAGGCATGAGCCAGCA 8132


RESULT 50
US-09-949-002-575/c
; Sequence 575, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 575
;   LENGTH: 75295
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(75295)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-002-575

  Query Match            55.2%;  Score 16;  DB 3;  Length 75295;
  Best Local Similarity  100.0%;  Pred. No. 12;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db      24181 CAGGCATGAGCCAGCA 24166


Search completed: April 12, 2006, 14:50:12
Job time : 150 secs
```

This page blank (uspto)

OM nucleic - nucleic search, using sw model

Run on: April 12, 2006, 13:45:08 ; Search time 3728 Seconds
(without alignments)
363.955 Million cell updates/sec

Title: SEQ1-4023-4051-4037A
Perfect score: 29
Sequence: 1 cctctctggagctcaggcatgagccagca 29

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 41078325 seqs, 23393541228 residues

Word size : 15

Total number of hits satisfying chosen parameters: 1409

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	19	65.5	557	11	FR0038584	AL126085 Fugu rubr
	2	19	65.5	565	5	C87641	C87641 C87641 Mous
	3	19	65.5	579	2	BG067834	BG067834 H3058F10-
c	4	19	65.5	949	11	CNS050CE	AL315383 Tetraodon
c	5	18	62.1	369	7	CV324980	CV324980 CM4-GN008
	6	18	62.1	530	5	BY706943	BY706943 BY706943
c	7	18	62.1	606	9	AQ500073	AQ500073 HS_5220_A
c	8	18	62.1	632	9	BZ896773	BZ896773 CH240_9N3
	9	18	62.1	747	6	CA466182	CA466182 AGENCOURT
	10	18	62.1	770	8	CX784900	CX784900 HESC3_40_
c	11	18	62.1	906	5	BQ676309	BQ676309 AGENCOURT
c	12	18	62.1	917	11	CT011342	CT011342 KBrH118L0
	13	18	62.1	931	9	CC298613	CC298613 CH261-178
	14	18	62.1	1055	9	CC221775	CC221775 CH261-92I
	15	18	62.1	1178	5	BU192345	BU192345 AGENCOURT
	16	17	58.6	260	8	DN895382	DN895382 nao66b10.
c	17	17	58.6	300	2	BB436030	BB436030 BB436030
	18	17	58.6	377	6	CF170455	CF170455 B0827H12-
	19	17	58.6	552	2	BE753088	BE753088 205980 MA
	20	17	58.6	611	3	BI889195	BI889195 ZF637-2-0
	21	17	58.6	676	6	CF169566	CF169566 B0815D08-
c	22	17	58.6	718	7	CV224516	CV224516 CS_hyp_16
c	23	17	58.6	802	10	AG490318	AG490318 Mus muscu
	24	17	58.6	809	9	AQ894819	AQ894819 HS_3133_A
	25	17	58.6	818	9	AZ752023	AZ752023 RPCI-24-1
c	26	17	58.6	918	9	CC475574	CC475574 CH240_301
	27	17	58.6	1104	9	CC189709	CC189709 CH261-131
	28	16	55.2	70	11	CR261805	CR261805 Forward s
	29	16	55.2	77	11	CR058210	CR058210 Forward s
	30	16	55.2	120	8	N88936	N88936 K6757F Huma
c	31	16	55.2	143	2	BF927034	BF927034 CM2-NT019
	32	16	55.2	150	2	BF800261	BF800261 CM4-CI006
c	33	16	55.2	168	2	BF882488	BF882488 CM1-ET019
	34	16	55.2	181	2	BB720345	BB720345 BB720345
c	35	16	55.2	191	10	CE664705	CE664705 tigr-gss-
c	36	16	55.2	222	9	B82265	B82265 RPCI11-14D2
c	37	16	55.2	239	9	AZ331080	AZ331080 1M0056B17
c	38	16	55.2	252	9	AZ800758	AZ800758 2M0058N22
	39	16	55.2	266	7	CV326948	CV326948 CM4-UT004
c	40	16	55.2	267	2	BI024787	BI024787 PM0-MT020
	41	16	55.2	270	1	AA579864	AA579864 nj41b10.s
c	42	16	55.2	273	9	AZ065220	AZ065220 RPCI-23-3
	43	16	55.2	282	9	AQ424457	AQ424457 CITBI-E1-
c	44	16	55.2	290	9	AQ052047	AQ052047 RPCI11-53
c	45	16	55.2	298	9	B79172	B79172 CIT978SK-11
c	46	16	55.2	298	9	B79288	B79288 CIT978SK-11
	47	16	55.2	309	2	BB400771	BB400771 BB400771
	48	16	55.2	311	11	CR069381	CR069381 Forward s
c	49	16	55.2	319	2	BG546523	BG546523 602574272
	50	16	55.2	320	1	AI445839	AI445839 tj12d08.x
c	51	16	55.2	322	1	AA082808	AA082808 zn25a01.r
	52	16	55.2	322	2	BB458677	BB458677 BB458677
	53	16	55.2	324	7	CN715839	CN715839 E0714G10-
c	54	16	55.2	324	9	AQ585183	AQ585183 RPCI-11-4
	55	16	55.2	333	1	AA814878	AA814878 oc06d10.s
	56	16	55.2	333	1	AA555188	AA555188 nl16g09.s
c	57	16	55.2	336	2	BF857164	BF857164 RC5-FT019
	58	16	55.2	348	7	CN702480	CN702480 E0462H06-
c	59	16	55.2	350	9	AZ283247	AZ283247 RPCI-23-4
c	60	16	55.2	358	9	AQ551659	AQ551659 RPCI-11-4
	61	16	55.2	359	1	AA564809	AA564809 nj36g05.s
	62	16	55.2	362	1	AA485487	AA485487 ab09c07.r
c	63	16	55.2	363	7	CR521575	CR521575 CR521575
c	64	16	55.2	363	9	AQ355538	AQ355538 CITBI-E1-
c	65	16	55.2	369	9	AQ005189	AQ005189 CIT-HSP-2
c	66	16	55.2	372	7	CV324003	CV324003 CM4-CT048
	67	16	55.2	373	1	AV646485	AV646485 AV646485
c	68	16	55.2	378	7	CV384605	CV384605 QV0-ST029
c	69	16	55.2	380	7	CV345282	CV345282 MR1-MT028
	70	16	55.2	382	5	BX953467	BX953467 DKFZp781M
c	71	16	55.2	386	10	AG202706	AG202706 Pan trogl
	72	16	55.2	387	2	BI188794	BI188794 d3b12fs.f
	73	16	55.2	389	8	H93143	H93143 yt91d04.s1
	74	16	55.2	390	1	AW301491	AW301491 xs77g10.x
	75	16	55.2	394	2	BF435882	BF435882 nab51f05.
c	76	16	55.2	401	8	H26293	H26293 yl51d01.s1
	77	16	55.2	401	9	AQ886333	AQ886333 HS_5542_B
	78	16	55.2	402	1	AW240228	AW240228 up30c02.y
	79	16	55.2	403	1	AI719047	AI719047 as65b09.x
c	80	16	55.2	403	9	AQ820449	AQ820449 HS_5323_B
	81	16	55.2	409	9	B68198	B68198 CIT978SK-A-
	82	16	55.2	421	1	AI049945	AI049945 an34h02.x
	83	16	55.2	423	1	AI547861	AI547861 UI-R-C3-s
	84	16	55.2	423	8	T06576	T06576 EST04465 Fe
	85	16	55.2	423	9	B70666	B70666 CIT-HSP-206
	86	16	55.2	425	1	AI866268	AI866268 wl28e09.x
	87	16	55.2	427	2	BB810089	BB810089 BB810089
c	88	16	55.2	428	1	AA233635	AA233635 zr43e07.r
c	89	16	55.2	432	1	AA809179	AA809179 nv38c12.r
	90	16	55.2	436	8	CX214307	CX214307 MNS22565
c	91	16	55.2	437	2	BF542555	BF542555 UI-R-C3-s
	92	16	55.2	437	2	BF547078	BF547078 UI-R-C1-k
c	93	16	55.2	437	2	BE982602	BE982602 UI-M-CG0p
c	94	16	55.2	437	9	AQ015439	AQ015439 CIT-HSP-2
c	95	16	55.2	438	9	AQ815609	AQ815609 HS_5348_A

c	96	16	55.2	439	9	AQ231147 AQ231147 HS_2026_B
	97	16	55.2	441	5	BY290699 BY290699 BY290699
c	98	16	55.2	443	8	H59956 H59956 yr16d05.r1
c	99	16	55.2	444	2	BE982887 BE982887 UI-M-CG0p
c	100	16	55.2	445	9	AZ500583 AZ500583 1M0339H02
	101	16	55.2	447	8	W88997 W88997 zh70h01.s1
	102	16	55.2	450	2	BF771599 BF771599 RC4-IT001
	103	16	55.2	451	2	BF731246 BF731246 mab83c06.
c	104	16	55.2	453	2	BI290861 BI290861 UI-R-DK0-
	105	16	55.2	453	5	BY250805 BY250805 BY250805
c	106	16	55.2	454	9	AQ369174 AQ369174 HS_5032_A
c	107	16	55.2	455	1	AW579146 AW579146 QV3-DT001
	108	16	55.2	457	8	R07499 R07499 ye97f04.s1
	109	16	55.2	459	1	AA011673 AA011673 zi03d06.s
c	110	16	55.2	459	9	AQ401423 AQ401423 HS_5054_A
	111	16	55.2	460	2	BF837925 BF837925 MR0-HT092
	112	16	55.2	461	2	BE160197 BE160197 QV1-HT041
	113	16	55.2	461	5	BY249093 BY249093 BY249093
	114	16	55.2	463	8	H71788 H71788 ys05e11.s1
	115	16	55.2	464	1	AI465694 AI465694 vw83b12.y
c	116	16	55.2	465	6	CB418072 CB418072 590836 MA
c	117	16	55.2	465	9	AZ443829 AZ443829 1M0238I15
	118	16	55.2	468	7	CV348569 CV348569 MR2-HN003
c	119	16	55.2	468	9	AQ121062 AQ121062 HS_3073_A
	120	16	55.2	469	1	AW073385 AW073385 xa39d12.x
	121	16	55.2	471	1	AA517859 AA517859 vh82e04.r
c	122	16	55.2	471	2	BG381447 BG381447 UI-R-CT0-
	123	16	55.2	474	1	AI447088 AI447088 mq90d01.x
c	124	16	55.2	476	1	AA446973 AA446973 zw90a01.r
	125	16	55.2	477	6	CA534769 CA534769 C0209E01-
c	126	16	55.2	479	9	AQ824037 AQ824037 HS_3141_A
	127	16	55.2	486	1	AW026687 AW026687 wv15g02.x
	128	16	55.2	486	1	AA527121 AA527121 ni26g07.s
	129	16	55.2	492	1	AW377859 AW377859 MR0-HT020
c	130	16	55.2	492	9	AQ354069 AQ354069 CITBI-E1-
c	131	16	55.2	493	2	BE586743 BE586743 WHE0514_H
c	132	16	55.2	494	2	BF908790 BF908790 QV1-UT009
	133	16	55.2	494	3	BM938797 BM938797 UI-M-CG0p
	134	16	55.2	500	8	N51443 N51443 yz17g02.s1
	135	16	55.2	503	1	AA028451 AA028451 mi19f12.r
	136	16	55.2	506	6	CD773385 CD773385 AGENCOURT
	137	16	55.2	509	6	CA559201 CA559201 K0255F04-
	138	16	55.2	510	10	AG200700 AG200700 Pan trogl
c	139	16	55.2	511	9	AQ233384 AQ233384 HS_2052_A
c	140	16	55.2	513	9	AQ756780 AQ756780 HS_5407_A
c	141	16	55.2	516	1	AL861876 AL861876 AL861876
c	142	16	55.2	517	1	AI648015 AI648015 uk39b10.x
	143	16	55.2	519	2	BB867531 BB867531 BB867531
c	144	16	55.2	519	9	AQ388228 AQ388228 RPCI11-14
c	145	16	55.2	519	9	AQ703076 AQ703076 HS_5466_A
c	146	16	55.2	520	9	AQ590262 AQ590262 HS_5399_A
	147	16	55.2	526	2	BF011910 BF011910 us42h07.y
c	148	16	55.2	528	7	CR630976 CR630976 DKFZp469A
	149	16	55.2	530	2	BG483390 BG483390 602504261
	150	16	55.2	532	1	AW377862 AW377862 MR0-HT020
c	151	16	55.2	532	9	BZ457918 BZ457918 BONEL64TR
c	152	16	55.2	534	6	CD239538 CD239538 FNPBVF10
c	153	16	55.2	536	1	AA724468 AA724468 ah99g02.s
	154	16	55.2	538	1	AA527198 AA527198 ni27a06.s
c	155	16	55.2	538	7	CR539706 CR539706 DKFZp459O
	156	16	55.2	540	9	AQ529492 AQ529492 RPCI-11-3
c	157	16	55.2	544	9	AZ788595 AZ788595 2M0035A22
c	158	16	55.2	545	9	AQ428708 AQ428708 CITBI-E1-
c	159	16	55.2	548	1	AA551030 AA551030 nk74a08.s
	160	16	55.2	550	2	BE145809 BE145809 MR0-HT020
c	161	16	55.2	552	2	BF970482 BF970482 602272647
	162	16	55.2	554	6	CD723153 CD723153 oj18e11.y
	163	16	55.2	555	6	CA536720 CA536720 C0237H05-
c	164	16	55.2	556	2	BE107275 BE107275 UI-R-BS1-
	165	16	55.2	556	6	CA559322 CA559322 K0257B09-
	166	16	55.2	557	6	CA560239 CA560239 K0269F05-
c	167	16	55.2	557	7	CR786618 CR786618 DKFZp468P
c	168	16	55.2	560	9	AQ262547 AQ262547 CITBI-E1-
c	169	16	55.2	562	9	AZ850864 AZ850864 2M0152E20
	170	16	55.2	565	5	BU783855 BU783855 in10f02.x
	171	16	55.2	567	3	BM875334 BM875334 ij54h07.x
c	172	16	55.2	570	5	BU951104 BU951104 io75h07.y
c	173	16	55.2	571	9	AQ668507 AQ668507 HS_5422_A
c	174	16	55.2	576	9	BH051631 BH051631 RPCI-24-3
c	175	16	55.2	578	6	CA590222 CA590222 hab23e06.
	176	16	55.2	584	6	CF749600 CF749600 UI-M-HJ0-
	177	16	55.2	584	9	AZ269841 AZ269841 RPCI-23-1
	178	16	55.2	584	9	AQ194932 AQ194932 RPCI11-47
	179	16	55.2	585	5	BU948134 BU948134 io50g07.x
	180	16	55.2	587	1	AJ273975 AJ273975 AJ273975
c	181	16	55.2	589	5	BX882903 BX882903 BX882903
	182	16	55.2	589	6	CD710684 CD710684 EST27211
c	183	16	55.2	590	5	BU661172 BU661172 cl68g12.z
c	184	16	55.2	590	5	BU675248 BU675248 UI-CF-DU1
c	185	16	55.2	591	6	CA951018 CA951018 ir93g07.y
c	186	16	55.2	591	9	AQ553311 AQ553311 RPCI-11-4
	187	16	55.2	593	7	CF916287 CF916287 B0993A05-
c	188	16	55.2	595	9	AZ817532 AZ817532 2M0087K04
c	189	16	55.2	596	9	AQ775163 AQ775163 HS_3151_A
c	190	16	55.2	596	9	AQ878403 AQ878403 HS_3069_A
c	191	16	55.2	599	2	BE143603 BE143603 MR0-HT016
	192	16	55.2	600	2	BG080512 BG080512 H3054E08-
	193	16	55.2	600	5	BX471088 BX471088 DKFZp686F
	194	16	55.2	602	1	AA523968 AA523968 ng28g03.s
c	195	16	55.2	604	6	CA944638 CA944638 UI-CF-FN0
	196	16	55.2	604	6	CF102147 CF102147 hac41h05.
c	197	16	55.2	604	9	BH744983 BH744983 gu74d08.b
	198	16	55.2	608	2	BI014292 BI014292 MR4-ET014
c	199	16	55.2	609	6	CF176856 CF176856 805275 MA
c	200	16	55.2	610	1	AL698554 AL698554 DKFZp686E
	201	16	55.2	611	6	CD564290 CD564290 B0476F08-
c	202	16	55.2	615	6	CA777023 CA777023 ip02a12.y
c	203	16	55.2	617	1	AL846599 AL846599 AL846599
	204	16	55.2	617	9	BZ891305 BZ891305 CH240_222
c	205	16	55.2	618	1	AW812376 AW812376 CM3-ST018
c	206	16	55.2	618	9	AQ442560 AQ442560 HS_5067_B
c	207	16	55.2	619	4	BC010386 BC010386 Homo sapi
	208	16	55.2	619	6	CF174294 CF174294 B0936H12-
	209	16	55.2	621	2	BG703313 BG703313 602685244
	210	16	55.2	623	5	BU579783 BU579783 im90g11.x
	211	16	55.2	624	7	CN715824 CN715824 E0714F03-
c	212	16	55.2	624	7	CR549735 CR549735 DKFZp469F
	213	16	55.2	625	3	BQ128197 BQ128197 ij79a08.x
c	214	16	55.2	626	5	BQ388992 BQ388992 NISC_mq05
	215	16	55.2	627	7	CN676675 CN676675 A0984B12-
	216	16	55.2	627	8	CX768038 CX768038 UI-M-HU0-
	217	16	55.2	631	10	AG090472 AG090472 Pan trogl
c	218	16	55.2	631	11	DE088469 DE088469 Oryzias l
c	219	16	55.2	639	1	AL849954 AL849954 AL849954
c	220	16	55.2	639	5	BX508998 BX508998 DKFZp686N
c	221	16	55.2	640	2	BB666275 BB666275 BB666275
c	222	16	55.2	640	5	BX269476 BX269476 BX269476
	223	16	55.2	640	6	CB268883 CB268883 1007790 H
c	224	16	55.2	641	9	AZ501580 AZ501580 1M0340C02
	225	16	55.2	642	5	BY743989 BY743989 BY743989
c	226	16	55.2	642	9	BZ923521 BZ923521 CH240_116
	227	16	55.2	643	3	BJ519965 BJ519965 BJ519965
	228	16	55.2	646	5	BU707824 BU707824 UI-M-FR0-
	229	16	55.2	646	5	BU783927 BU783927 in11f09.x
	230	16	55.2	646	7	CN267841 CN267841 170005318
	231	16	55.2	648	1	AV687842 AV687842 AV687842
	232	16	55.2	649	9	BH279586 BH279586 CH230-191
c	233	16	55.2	650	6	CD595984 CD595984 RK099A4B1
c	234	16	55.2	650	7	CR545367 CR545367 DKFZp459L
c	235	16	55.2	651	10	AG074773 AG074773 Pan trogl
	236	16	55.2	651	10	AG137057 AG137057 Pan trogl
c	237	16	55.2	652	11	CR233493 CR233493 Reverse s
	238	16	55.2	654	10	CE292644 CE292644 tigr-gss-
	239	16	55.2	654	10	CL696867 SP__Ba000
	240	16	55.2	655	9	BH275891 BH275891 CH230-70J
c	241	16	55.2	658	10	AG037709 AG037709 Pan trogl

c	242	16	55.2	659	3	BJ066019 BJ066019 BJ066019
	243	16	55.2	660	3	BM944858 UI-M-EH0p BM944858
c	244	16	55.2	661	10	AG090459 AG090459 Pan trogl
	245	16	55.2	662	5	BY739543 BY739543 BY739543
c	246	16	55.2	664	10	AG140854 AG140854 Pan trogl
	247	16	55.2	665	5	BQ446428 BQ446428 UI-H-EU1-
	248	16	55.2	666	7	CF948670 CF948670 UI-M-HJ0-
	249	16	55.2	667	2	BB398773 BB398773 BB398773
	250	16	55.2	670	2	BG482679 BG482679 602502639
c	251	16	55.2	673	11	CR223837 CR223837 Reverse s
	252	16	55.2	677	5	BY733283 BY733283 BY733283
	253	16	55.2	677	8	DN999150 DN999150 TC100254
	254	16	55.2	683	10	AG174852 AG174852 Pan trogl
	255	16	55.2	684	10	AG130748 AG130748 Pan trogl
c	256	16	55.2	688	10	AG088518 AG088518 Pan trogl
	257	16	55.2	689	3	BJ877032 BJ877032 BJ877032
c	258	16	55.2	689	10	AG049129 AG049129 Pan trogl
	259	16	55.2	691	2	BF797984 BF797984 602258716
c	260	16	55.2	691	10	AG094253 AG094253 Pan trogl
c	261	16	55.2	694	10	AG185790 AG185790 Pan trogl
	262	16	55.2	697	2	BB642940 BB642940 BB642940
c	263	16	55.2	697	10	AG164198 AG164198 Pan trogl
c	264	16	55.2	704	2	BE261995 BE261995 601152668
	265	16	55.2	704	10	AG123672 AG123672 Pan trogl
c	266	16	55.2	705	7	CJ032952 CJ032952 CJ032952
	267	16	55.2	706	2	BB295050 BB295050 BB295050
c	268	16	55.2	706	9	BH956400 BH956400 odi92g04.
c	269	16	55.2	709	5	BX508296 BX508296 DKFZp686M
c	270	16	55.2	712	2	BI156848 BI156848 602921483
c	271	16	55.2	715	9	AQ321056 AQ321056 RPCI11-10
	272	16	55.2	716	9	BH955355 BH955355 odg91e12.
	273	16	55.2	717	7	CR787988 CR787988 DKFZp459K
c	274	16	55.2	721	3	BP140405 BP140405 BP140405
	275	16	55.2	723	10	AG117018 AG117018 Pan trogl
c	276	16	55.2	733	2	BG709541 BG709541 602673547
	277	16	55.2	737	2	BG173082 BG173082 602336652
	278	16	55.2	737	3	BI411383 BI411383 602964746
	279	16	55.2	737	3	BJ737859 BJ737859 BJ737859
	280	16	55.2	738	10	BX177798 BX177798 Danio rer
c	281	16	55.2	738	11	CR188193 CR188193 Reverse s
	282	16	55.2	739	2	BG404271 BG404271 602420344
	283	16	55.2	747	7	CN535277 CN535277 UI-M-HS0-
	284	16	55.2	748	8	CX757107 CX757107 AGENCOURT
	285	16	55.2	749	2	BG548705 BG548705 602576316
	286	16	55.2	751	2	BG492457 BG492457 602536403
	287	16	55.2	751	9	BH202059 BH202059 Sm1-62J22
c	288	16	55.2	752	10	AG468836 AG468836 Mus muscu
c	289	16	55.2	754	11	CR153956 CR153956 Reverse s
	290	16	55.2	757	3	BM047913 BM047913 603618974
c	291	16	55.2	760	9	BZ771187 BZ771187 mcs73c09.
c	292	16	55.2	760	10	AG168412 AG168412 Pan trogl
c	293	16	55.2	764	9	BH319455 BH319455 CH230-79I
c	294	16	55.2	771	9	BZ608539 BZ608539 WHACA91TF
c	295	16	55.2	771	10	AG599106 AG599106 Mus muscu
	296	16	55.2	780	3	BI559363 BI559363 603253033
	297	16	55.2	782	5	BX373196 BX373196 BX373196
	298	16	55.2	784	2	BI090422 BI090422 602853706
c	299	16	55.2	785	5	BX116744 BX116744 BX116744
	300	16	55.2	790	7	CV112311 CV112311 AGENCOURT
	301	16	55.2	791	2	BG976730 BG976730 602847073
	302	16	55.2	791	2	BI161338 BI161338 602865736
c	303	16	55.2	797	3	BP436526 BP436526 BP436526
c	304	16	55.2	798	9	BH580557 BH580557 BOGXW15TF
	305	16	55.2	801	5	BX495074 BX495074 DKFZp779D
c	306	16	55.2	802	6	CD518084 CD518084 AGENCOURT
	307	16	55.2	804	7	CK793570 CK793570 AGENCOURT
c	308	16	55.2	806	5	BU103831 BU103831 603006471
	309	16	55.2	809	1	AU117791 AU117791 AU117791
	310	16	55.2	809	10	BX195591 BX195591 Danio rer
	311	16	55.2	811	10	AG475940 AG475940 Mus muscu
	312	16	55.2	811	10	BX153056 BX153056 Danio rer
c	313	16	55.2	814	5	BU567680 BU567680 AGENCOURT
	314	16	55.2	816	7	CK795304 CK795304 AGENCOURT
c	315	16	55.2	822	4	CNS0FLKR CR677311 Tetraodon
c	316	16	55.2	823	4	CNS0FB2E CR663680 Tetraodon
	317	16	55.2	827	5	BU613095 BU613095 UI-M-FR0-
	318	16	55.2	829	8	CX242509 CX242509 NMA00288
c	319	16	55.2	830	5	BX643900 BX643900 DKFZp781C
c	320	16	55.2	831	2	BF674597 BF674597 602137815
	321	16	55.2	833	10	CW716542 CW716542 AIAA-aab7
c	322	16	55.2	834	5	BU115811 BU115811 603140350
	323	16	55.2	835	8	CX762002 CX762002 AGENCOURT
	324	16	55.2	838	5	BU856501 BU856501 AGENCOURT
c	325	16	55.2	839	9	BH669066 BH669066 BOMNO54TR
c	326	16	55.2	845	2	BG619789 BG619789 602619353
c	327	16	55.2	846	2	BI156046 BI156046 602903601
	328	16	55.2	847	9	CC520975 CC520975 CH240_368
	329	16	55.2	847	10	CW871112 CW871112 shS2kd39-
	330	16	55.2	850	9	BZ457197 BZ457197 BONPF60TR
	331	16	55.2	856	2	BG569986 BG569986 602590240
c	332	16	55.2	858	2	BG393677 BG393677 602412095
	333	16	55.2	858	8	CX239438 CX239438 NMA07281
c	334	16	55.2	862	1	AU136517 AU136517 AU136517
	335	16	55.2	866	2	BG864508 BG864508 602798531
	336	16	55.2	869	2	BG334472 BG334472 602461105
c	337	16	55.2	869	5	BQ428123 BQ428123 AGENCOURT
	338	16	55.2	870	11	CR213495 CR213495 Reverse s
c	339	16	55.2	878	8	CX981773 CX981773 JGI_CAAP1
c	340	16	55.2	880	5	BU904367 BU904367 AGENCOURT
c	341	16	55.2	880	11	CNS04FMB AL288524 Tetraodon
	342	16	55.2	886	3	BI731877 BI731877 603353621
c	343	16	55.2	890	11	CR087446 CR087446 Reverse s
c	344	16	55.2	891	9	AQ243819 AQ243819 HS_2062_B
c	345	16	55.2	895	2	BF679271 BF679271 602153463
	346	16	55.2	900	5	BQ964001 BQ964001 AGENCOURT
c	347	16	55.2	901	2	BF686049 BF686049 602142672
	348	16	55.2	901	4	AK045109 AK045109 Mus muscu
c	349	16	55.2	905	5	BU905717 BU905717 AGENCOURT
c	350	16	55.2	910	4	CNS0GHUF CR719153 Tetraodon
c	351	16	55.2	912	6	CD559148 CD559148 AGENCOURT
c	352	16	55.2	917	10	CNS02ZMP AL221146 Tetraodon
	353	16	55.2	926	2	BE731053 BE731053 601570129
c	354	16	55.2	932	9	BZ598051 BZ598051 WHADQ15TF
c	355	16	55.2	943	9	CC202026 CC202026 CH261-164
c	356	16	55.2	947	4	CNS0GF96 CR715790 Tetraodon
	357	16	55.2	950	5	BQ886125 BQ886125 AGENCOURT
c	358	16	55.2	964	9	BZ164211 BZ164211 CH230-246
	359	16	55.2	987	7	CK295388 CK295388 EST758102
	360	16	55.2	994	5	BU511177 BU511177 AGENCOURT
c	361	16	55.2	998	9	BZ602381 BZ602381 WHACR41TR
	362	16	55.2	1008	10	AY417434 AY417434 Mus muscu
	363	16	55.2	1036	5	BY712561 BY712561 BY712561
	364	16	55.2	1044	2	BE867692 BE867692 601443247
	365	16	55.2	1074	2	BG696469 BG696469 602659579
c	366	16	55.2	1857	4	CR610740 CR610740 full-leng
	367	16	55.2	1905	4	AK013599 AK013599 Mus muscu
	368	16	55.2	2516	4	AK049174 AK049174 Mus muscu
	369	16	55.2	2727	4	BC026509 BC026509 Mus muscu
	370	16	55.2	2860	4	AK079317 AK079317 Mus muscu
c	371	16	55.2	3895	4	CR860742 CR860742 Pongo pyg
c	372	16	55.2	4023	4	AK090391 AK090391 Mus muscu
	373	16	55.2	5920	4	AK029876 AK029876 Mus muscu
c	374	15	51.7	105	6	CB948503 CB948503 AGENCOURT
c	375	15	51.7	108	7	CV390855 CV390855 QV2-HT069
c	376	15	51.7	112	8	H13973 H13973 EST00065 Ch
	377	15	51.7	119	1	AW626016 AW626016 EST319923
c	378	15	51.7	134	1	AA250272 AA250272 mz61b01.r
c	379	15	51.7	140	1	AA242686 AA242686 mx08h02.r
c	380	15	51.7	142	2	BG002183 BG002183 MR3-GN018
c	381	15	51.7	144	1	AA255048 AA255048 mz85f10.r
	382	15	51.7	153	10	CE816057 CE816057 tigr-gss-
c	383	15	51.7	154	1	AA245016 AA245016 mx09d07.r
	384	15	51.7	164	1	AA873663 AA873663 oe02h02.s
	385	15	51.7	170	10	CE515628 CE515628 tigr-gss-
	386	15	51.7	175	1	AW110751 AW110751 MT1548 mo
c	387	15	51.7	178	9	BH198183 BH198183 TC3-67G16

	388	15	51.7	181	1	AW110668 AW110668 MT1248 mo
	389	15	51.7	184	1	AW110687 AW110687 MT1848 mo
c	390	15	51.7	187	2	BF959662 BF959662 RC5-NN106
c	391	15	51.7	188	1	AW799517 AW799517 PM2-UM005
	392	15	51.7	195	1	AW110693 AW110693 MT808 mou
	393	15	51.7	197	2	BE486188 BE486188 173764 BA
c	394	15	51.7	201	9	AQ196254 AQ196254 RPCI11-67
c	395	15	51.7	202	1	AA659774 AA659774 nu84h09.s
c	396	15	51.7	205	1	AV009050 AV009050 AV009050
	397	15	51.7	209	1	AI314642 AI314642 uj26h08.x
c	398	15	51.7	209	10	CE643967 CE643967 tigr-gss-
c	399	15	51.7	214	1	AV017356 AV017356 AV017356
c	400	15	51.7	217	9	CC436282 CC436282 PUHCH20TD
c	401	15	51.7	217	10	CE648501 CE648501 tigr-gss-
	402	15	51.7	218	1	AW897539 AW897539 CM0-NN005
c	403	15	51.7	219	1	AV018975 AV018975 AV018975
c	404	15	51.7	219	9	AZ123720 AZ123720 RPCI-23-5
	405	15	51.7	226	11	DR23B21T AL981366 Danio rer
c	406	15	51.7	229	2	BF595566 BF595566 su66f10.y
	407	15	51.7	231	2	BF821162 BF821162 RC4-RT005
c	408	15	51.7	232	1	AI468205 AI468205 tg56e08.x
c	409	15	51.7	233	1	AI814200 AI814200 wk54e12.x
c	410	15	51.7	235	1	AV028297 AV028297 AV028297
c	411	15	51.7	236	1	AV028052 AV028052 AV028052
c	412	15	51.7	236	9	AZ902895 AZ902895 RPCI-24-1
c	413	15	51.7	242	2	BF840335 BF840335 RC3-HT023
	414	15	51.7	245	1	AW110685 AW110685 MT1213 mo
c	415	15	51.7	245	8	CX750654 CX750654 JGI_ANHP2
c	416	15	51.7	247	1	AW602920 AW602920 CM0-BT060
c	417	15	51.7	249	3	BQ027750 BQ027750 UI-H-CO0-
	418	15	51.7	250	2	BE151356 BE151356 CM2-HT028
c	419	15	51.7	250	6	CB947275 CB947275 AGENCOURT
c	420	15	51.7	254	1	AA488958 AA488958 aa55f02.r
	421	15	51.7	255	1	AU042969 AU042969 AU042969
	422	15	51.7	255	9	AQ274801 AQ274801 RPCI-6-13
	423	15	51.7	256	3	BM429842 BM429842 1Duo22A4.
	424	15	51.7	259	8	CX740517 CX740517 JGI_XZF16
c	425	15	51.7	261	1	AV026905 AV026905 AV026905
c	426	15	51.7	263	1	AA669085 AA669085 aa81e01.s
	427	15	51.7	264	1	AU040330 AU040330 AU040330
c	428	15	51.7	269	8	Z35033 Z35033 ATTS3672 Gi
c	429	15	51.7	270	8	CV981057 CV981057 UMC-bemiv
	430	15	51.7	273	2	BF751634 BF751634 MR0-BN011
	431	15	51.7	275	10	BX129485 BX129485 Danio rer
c	432	15	51.7	276	9	AZ905907 AZ905907 RPCI-24-1
c	433	15	51.7	278	9	AZ903279 AZ903279 RPCI-24-1
	434	15	51.7	279	1	AW337710 AW337710 wu99c02.x
c	435	15	51.7	280	10	BX125138 BX125138 Danio rer
c	436	15	51.7	281	2	BI006965 BI006965 RC5-RT005
c	437	15	51.7	281	7	CV677695 CV677695 ih38d03.k
	438	15	51.7	282	10	BX205671 BX205671 Danio rer
c	439	15	51.7	284	1	AV024997 AV024997 AV024997
c	440	15	51.7	285	1	AA619953 AA619953 vl59b04.s
	441	15	51.7	286	10	CL278336 CL278336 Ggal_102a
c	442	15	51.7	287	2	BB717062 BB717062 BB717062
	443	15	51.7	288	1	AW415395 AW415395 49658 MAR
c	444	15	51.7	288	7	CK730003 CK730003 UMC-bov_0
c	445	15	51.7	289	1	AV012275 AV012275 AV012275
c	446	15	51.7	290	1	AW292687 AW292687 UI-H-BI2-
	447	15	51.7	297	1	AV097552 AV097552 AV097552
	448	15	51.7	300	1	AI846790 AI846790 UI-M-AO1-
c	449	15	51.7	300	2	BF935355 BF935355 CM1-NT024
	450	15	51.7	301	6	CB700840 CB700840 AMGNNUC:M
	451	15	51.7	304	9	AQ722579 AQ722579 HS_5236_A
	452	15	51.7	305	1	AW110695 AW110695 MT5497 mo
c	453	15	51.7	306	1	AW603852 AW603852 CM0-CN004
c	454	15	51.7	307	1	AA291101 AA291101 zs44g12.r
c	455	15	51.7	307	6	CD313385 CD313385 StrPu621.
c	456	15	51.7	309	2	BF927407 BF927407 CM2-NT019
c	457	15	51.7	310	2	BG979463 BG979463 CM4-CN006
c	458	15	51.7	310	9	AZ293437 AZ293437 RPCI-23-1
	459	15	51.7	315	1	AW110683 AW110683 MT1561 mo
c	460	15	51.7	316	1	AA578774 AA578774 nh24a05.s
c	461	15	51.7	317	6	CD580681 CD580681 RK005A2H0
	462	15	51.7	320	7	CR521297 CR521297 CR521297
	463	15	51.7	322	8	CX750655 CX750655 JGI_ANHP2
c	464	15	51.7	322	9	B75288 B75288 RPCI11-15B9
	465	15	51.7	323	1	AW113299 AW113299 MC887 mou
	466	15	51.7	325	2	BF807409 BF807409 RC2-CI009
	467	15	51.7	325	10	BX221154 BX221154 Danio rer
	468	15	51.7	326	1	AL643684 AL643684 AL643684
c	469	15	51.7	327	1	AA386301 AA386301 EST185044
	470	15	51.7	331	1	AI869760 AI869760 wm03b10.x
c	471	15	51.7	331	9	CC550354 CC550354 CH240_435
	472	15	51.7	334	2	BI007494 BI007494 MR1-RT007
c	473	15	51.7	336	1	AA837227 AA837227 od69c10.s
c	474	15	51.7	336	5	BY190819 BY190819 BY190819
c	475	15	51.7	336	6	CB948609 CB948609 AGENCOURT
c	476	15	51.7	341	5	BY314483 BY314483 BY314483
	477	15	51.7	342	1	AW885804 AW885804 RC4-OT007
	478	15	51.7	342	10	BX160032 BX160032 Danio rer
	479	15	51.7	343	10	BX159604 BX159604 Danio rer
	480	15	51.7	344	9	AQ921590 AQ921590 RPCI-23-2
c	481	15	51.7	345	5	BY186892 BY186892 BY186892
c	482	15	51.7	345	6	CB948085 CB948085 AGENCOURT
	483	15	51.7	346	1	AI415011 AI415011 mb72a12.x
c	484	15	51.7	346	1	AV828281 AV828281 AV828281
c	485	15	51.7	346	10	CE703903 CE703903 tigr-gss-
	486	15	51.7	349	1	AW110686 AW110686 MT2051 mo
c	487	15	51.7	349	2	BF922895 BF922895 QV4-NT024
	488	15	51.7	352	1	AW113298 AW113298 MC362 mou
	489	15	51.7	353	1	AW110696 AW110696 MT1702 mo
	490	15	51.7	353	9	AQ198457 AQ198457 RPCI11-62
c	491	15	51.7	354	1	AA087039 AA087039 mk20d07.r
	492	15	51.7	354	1	AW110691 AW110691 MT1339 mo
	493	15	51.7	354	1	AW110697 AW110697 MT1127 mo
	494	15	51.7	355	1	AW110669 AW110669 MT17475 m
	495	15	51.7	357	2	BE069060 BE069060 QV3-BT037
c	496	15	51.7	358	3	BP807499 BP807499 BP807499
c	497	15	51.7	360	1	AI057469 AI057469 ow80e01.x
	498	15	51.7	360	3	BM030929 BM030929 495481 MA
c	499	15	51.7	363	8	CX698499 CX698499 yde27b08.
c	500	15	51.7	365	1	AV008009 AV008009 AV008009

ALIGNMENTS

RESULT 1
FR0038584

LOCUS	FR0038584 557 bp DNA linear GSS 25-FEB-2004
DEFINITION	Fugu rubripes GSS sequence, clone 076D19aE11, genomic survey sequence.
ACCESSION	AL126085
VERSION	AL126085.1 GI:6107700
KEYWORDS	GSS; genome survey sequence.
SOURCE	Takifugu rubripes (Fugu rubripes)
ORGANISM	Takifugu rubripes Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetradontoidea; Tetraodontidae; Takifugu.
REFERENCE	1
AUTHORS	Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J., Bouchireb,N., Cottage,A., Yeo,G.S., Umrania,Y., Williams,G. and Brenner,S.
TITLE	Generation and analysis of 25 Mb of genomic DNA from the pufferfish Fugu rubripes by sequence scanning
JOURNAL	Genome Res. 9 (10), 960-971 (1999)
PUBMED	10523524
REFERENCE	2 (bases 1 to 557)
AUTHORS	Elgar,G., Clark,M.S., Smith,S., Meek,S., Warner,S., Edwards,Y.J.K., Umrania,Y., Williams,G. and Brenner,S.
TITLE	Direct Submission
JOURNAL	Submitted (11-OCT-1999) MRC Human Genome Mapping Project Resource Centre, Hinxton, Cambridge, CB10 1SB. UK Email:

```
            biohelp@hgmp.mrc.ac.uk
COMMENT     Vector: pBluescript II KS
            V_type: phagemid
            PRIMER: KS
            DESCR:
            One pass dye-terminator sequencing of cosmid cloned genomic
            sequence.
FEATURES             Location/Qualifiers
     source          1. .557
                     /organism="Takifugu rubripes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:31033"
                     /clone="076D19aE11"
                     /clone_lib="cosmid 076D19"
ORIGIN

  Query Match             65.5%;  Score 19;  DB 11;  Length 557;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 CTGGAGCTCAGGCATGAGC 24
              |||||||||||||||||||
Db        105 CTGGAGCTCAGGCATGAGC 123


RESULT 2
C87641
LOCUS       C87641                   565 bp    mRNA    linear   EST 11-MAR-1998
DEFINITION  C87641 Mouse fertilized one-cell-embryo cDNA Mus musculus cDNA
            clone J0248F08 3', mRNA sequence.
ACCESSION   C87641
VERSION     C87641.1  GI:2919598
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 565)
  AUTHORS   Ko,M.S.H, Kitchen,J.R., Wang,X., Wang,X., Threat,T.A., Sun,T.,
            DePalma,G.E., Liang,Y., Kargul,G.J., Sharara,R., Paonessa,P.D. and
            Doi,H.
  TITLE     Systematic analyses of genes expressed in fertilized mouse eggs
            (The ERATO/Doi Project at Wayne State University)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Hirofumi Doi
            Doi Bioasymmetry Project, ERATO
            Japan Science and Technology Corporation (JST)
            WBG Marive East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan
            Email: hd@bioa.jst.go.jp.
FEATURES             Location/Qualifiers
     source          1. .565
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="J0248F08"
                     /dev_stage="fertilized one-cell-embryo"
                     /clone_lib="Mouse fertilized one-cell-embryo cDNA"
ORIGIN

  Query Match             65.5%;  Score 19;  DB 5;  Length 565;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAGCCA 26
              |||||||||||||||||||
Db        280 GGAGCTCAGGCATGAGCCA 298


RESULT 3
BG067834
LOCUS       BG067834                 579 bp    mRNA    linear   EST 17-DEC-2003
DEFINITION  H3058F10-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
            H3058F10 3', mRNA sequence.
ACCESSION   BG067834
VERSION     BG067834.2  GI:40015010
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 579)
  AUTHORS   Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X.,
            Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H.,
            Wood,W.H. III, Becker,K.G. and Ko,M.S.H.
  TITLE     Genome-wide expression profiling of mid-gestation placenta and
            embryo using a 15,000 mouse developmental cDNA microarray
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
  PUBMED    10922068
COMMENT     On Jan 26, 2001 this sequence version replaced gi:12550403.
            Other_ESTs: H3058F10-5
            Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            This clone set has been freely distributed to the community. Please
            visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
            Plate: H3058  row: F  column: 10
            Seq primer: -21M13 Forward
            High quality sequence stop: 579
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .579
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:H3058F10-3"
                     /db_xref="taxon:10090"
                     /clone="H3058F10"
                     /sex="Clones arrayed from a variety of cDNA libraries"
                     /dev_stage="Clones arrayed from a variety of cDNA
                     libraries"
                     /lab_host="DH10B"
                     /clone_lib="NIA Mouse 15K cDNA Clone Set"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
                     clone is among a rearrayed set of 15,247 clones from 11
                     embryo  cDNA libraries (including preimplantation stage
                     embryos from  unfertilized egg to blastocyst, embryonic
                     part of E7.5 embryos,  extraembryonic part of E7.5
                     embryos, and E12.5 female mesonephros/gonad)  and one
                     newborn ovary cDNA library. Average insert size 1.5 kb.
                     All  source libraries are cloned unidirectionally with
                     Oligo(dT)-Not primers.  References include: (1)
                     Genome-wide expression profiling of mid-gestation
                     placenta and embryo using a 15,000 mouse developmental
                     cDNA microarray,  2000, Proc. Natl. Acad. Sci. U S A, 97:
                     9127-9132; (2) Large-scale cDNA  analysis reveals phased
                     gene expression patterns during preimplantation  mouse
                     develolpment, 2000, Development, 127: 1737-1749; (3)
                     Genome-wide  mapping of unselected transcripts from
                     extraembryonic tissue of 7.5-day  mouse embryos reveals
                     enrichment in the t-complex and under-representation  on
                     the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."
ORIGIN

  Query Match             65.5%;  Score 19;  DB 2;  Length 579;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAGCCA 26
              |||||||||||||||||||
Db        317 GGAGCTCAGGCATGAGCCA 335
```

```
RESULT 4
CNS050CE/c
LOCUS       CNS050CE                949 bp    DNA     linear   GSS 26-JUL-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            029F10 of library B from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL315383
VERSION     AL315383.1  GI:9548271
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetradontoidea; Tetraodontidae; Tetraodon.
REFERENCE   1
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Estimate of human gene number provided by genome-wide analysis
            using Tetraodon nigroviridis DNA sequence
  JOURNAL   Nat. Genet. 25 (2), 235-238 (2000)
   PUBMED   10835645
REFERENCE   2
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
            Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
            Saurin,W., Bernot,A. and Weissenbach,J.
  TITLE     Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Genome Res. 10 (7), 939-949 (2000)
   PUBMED   10899143
REFERENCE   3  (bases 1 to 949)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .949
                     /organism="Tetraodon nigroviridis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:99883"
                     /clone="029F10"
                     /clone_lib="B"
                     /note="Genoscope sequence ID : C0AB029DC05C1
                     end : T7"
ORIGIN

  Query Match             65.5%;  Score 19;  DB 11;  Length 949;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAGGCA 19
              |||||||||||||||||||
Db        596 CCTCTCTGGAGCTCAGGCA 578


RESULT 5
CV324980/c
LOCUS       CV324980                369 bp    mRNA    linear   EST 24-SEP-2004
DEFINITION  CM4-GN0084-160900-315-e12 GN0084 Homo sapiens cDNA, mRNA sequence.
ACCESSION   CV324980
VERSION     CV324980.1  GI:52648194
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 369)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. http://www.ludwig.org.br.
FEATURES             Location/Qualifiers
     source          1. .369
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0084"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."
ORIGIN

  Query Match             62.1%;  Score 18;  DB 7;  Length 369;
  Best Local Similarity   100.0%;  Pred. No. 52;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAGCC 25
              ||||||||||||||||||
Db        307 GGAGCTCAGGCATGAGCC 290


RESULT 6
BY706943
LOCUS       BY706943                530 bp    mRNA    linear   EST 16-DEC-2002
DEFINITION  BY706943 RIKEN full-length enriched, adult male testis Mus musculus
            cDNA clone 1700071G13 5', mRNA sequence.
ACCESSION   BY706943
VERSION     BY706943.1  GI:27118113
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 530)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
```

Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

PUBMED 12466851

COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES Location/Qualifiers

source 1. .530
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1700071G13"
/sex="male"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="SOLR"
/clone_lib="RIKEN full-length enriched, adult male testis"
/note="Site_1: XhoI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGAGCGGCCGCAATTAATTCTCGAGTTAATTAAATTAATCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. "

ORIGIN

Query Match 62.1%; Score 18; DB 5; Length 530;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTCTCTGGAGCTCAGGC 18
||||||||||||||||||
Db 6 CCTCTCTGGAGCTCAGGC 23

RESULT 7
AQ500073/c
LOCUS AQ500073 606 bp DNA linear GSS 28-APR-1999
DEFINITION HS_5220_A2_F09_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=796 Col=18 Row=K, genomic survey sequence.
ACCESSION AQ500073
VERSION AQ500073.1 GI:4692646
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 606)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
Plate: 796 row: K column: 18
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 606.
FEATURES Location/Qualifiers
source 1. .606
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=796 Col=18 Row=K"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at EcoRI sites"
ORIGIN

Query Match 62.1%; Score 18; DB 9; Length 606;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         11 GCTCAGGCATGAGCCAGC 28
              ||||||||||||||||||
Db        313 GCTCAGGCATGAGCCAGC 296


RESULT 8
BZ896773/c
LOCUS       BZ896773                632 bp    DNA     linear   GSS 12-JUN-2003
DEFINITION  CH240_9N3.TV CHORI-240 Bos taurus genomic clone CH240_9N3, genomic
            survey sequence.
ACCESSION   BZ896773
VERSION     BZ896773.1  GI:31621824
KEYWORDS    GSS.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 632)
  AUTHORS   Larkin,D.M., Everts-van der Wind,A., Rebeiz,M., Schweitzer,P.A.,
            Bachman,S., Green,C., Wright,C.L., Campos,E.J., Benson,L.D.,
            Edwards,J., Liu,L., Osoegawa,K., Womack,J.E., de Jong,P.J. and
            Lewin,H.A.
  TITLE     A cattle-human comparative map built with cattle BAC-ends and human
            genome sequence
  JOURNAL   Genome Res. 13 (8), 1966-1972 (2003)
   PUBMED   12902387
COMMENT     Other_GSSs: CH240_9N3.TJ
            Contact: Harris Lewin
            Department of Animal Sciences
            University of Illinois at Urbana Champaign
            1201 W. Gregory Dr., Urbana, IL 61801, USA
            Tel: 217 333 5998
            Fax: 217 244 5617
            Email: h-lewin@uiuc.edu
            Clones are derived from the bovine BAC library CHORI-240
            (http://www.chori.org/bacpac/bovine240.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/ordering_information.htm). This work
            was undertaken as part of the International Bovine BAC Mapping
            Consortium (IBBMC) by by University of Illinois at Urbana
            Champaign, USA with funds provided by grant No. AG202-34480-11828
            from USDA-CSREES and AG99-35205-8534 from USDA/NRI (Livestock
            Genome Sequencing Initiative)
            Plate: 9  row: N  column: 3
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .632
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /strain="breed: Hereford"
                     /db_xref="taxon:9913"
                     /clone="CH240_9N3"
                     /sex="Male"
                     /cell_type="Blood"
                     /clone_lib="CHORI-240"
                     /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
                     Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
                     library (Male) produced by Pieter de Jong"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 9;  Length 632;
  Best Local Similarity   100.0%;  Pred. No. 54;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         12 CTCAGGCATGAGCCAGCA 29
              ||||||||||||||||||
Db        315 CTCAGGCATGAGCCAGCA 298


RESULT 9
CA466182
LOCUS       CA466182                747 bp    mRNA    linear   EST 12-NOV-2002
DEFINITION  AGENCOURT_10728941 NIH_MGC_169 Mus musculus cDNA clone
            IMAGE:6774514 5', mRNA sequence.
ACCESSION   CA466182
VERSION     CA466182.1  GI:24922534
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 747)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jonathan Kuo, NIMH
             cDNA Library Preparation: Michael Brownstein Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM3099  row: a  column: 09
            High quality sequence stop: 341.
FEATURES             Location/Qualifiers
     source          1. .747
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6774514"
                     /lab_host="DH10B (T1-phage-resistant)"
                     /clone_lib="NIH_MGC_169"
                     /note="Organ: Testicles; Vector: pDNR-LIB; Site_1: SfiI
                     (ggccattatggcc); Site_2: SfiI (ggccgcctcggcc); cDNA made
                     by oligo-dT priming and directionally cloned. 5' and 3'
                     adaptors were used in cloning as follows:
                     5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'. Full-length
                     enriched library was constructed using the Clontech
                     Creator SMART kit and size-selected to contain the 0.5 kb
                     size fraction. Library created in thelaboratory of M.
                     Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             62.1%;  Score 18;  DB 6;  Length 747;
  Best Local Similarity   100.0%;  Pred. No. 54;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAGGC 18
              ||||||||||||||||||
Db         37 CCTCTCTGGAGCTCAGGC 54


RESULT 10
CX784900
LOCUS       CX784900                770 bp    mRNA    linear   EST 02-MAR-2005
DEFINITION  HESC3_40_E01.g1_A036 NIH_MGC_260 Homo sapiens cDNA clone
            IMAGE:7479268 5', mRNA sequence.
ACCESSION   CX784900
VERSION     CX784900.1  GI:58301690
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 770)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15795  row: j  column: 02
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 770.
FEATURES             Location/Qualifiers
     source          1. .770
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7479268"
                     /sex="male"
                     /tissue_type="human embryonic stem cells"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_260"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos. Cell line
                     id and NIH Registry designation is BG01. Positive for
                     SSEA3, SSEA4, Tra 1-60, Tra 1-81, CD9, Alk Phos, Oct4 and
                     Nanog expression; negative for SSEA1 expression. Passage
                     number 21. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_261). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             62.1%;  Score 18;  DB 8;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 54;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 CTGGAGCTCAGGCATGAG 23
              ||||||||||||||||||
Db         13 CTGGAGCTCAGGCATGAG 30


RESULT 11
BQ676309/c
LOCUS       BQ676309                906 bp    mRNA    linear   EST 15-JUL-2002
DEFINITION  AGENCOURT_8210305 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6259178
            5', mRNA sequence.
ACCESSION   BQ676309
VERSION     BQ676309.1  GI:21788988
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 906)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2416  row: e  column: 03
            High quality sequence stop: 567.
FEATURES             Location/Qualifiers
     source          1. .906
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6259178"
                     /tissue_type="melanotic melanoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_112"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed  by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of  California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

  Query Match             62.1%;  Score 18;  DB 5;  Length 906;
  Best Local Similarity   100.0%;  Pred. No. 55;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 TGGAGCTCAGGCATGAGC 24
              ||||||||||||||||||
Db        639 TGGAGCTCAGGCATGAGC 622


RESULT 12
CT011342/c
LOCUS       CT011342                917 bp    DNA     linear   GSS 09-AUG-2005
DEFINITION  KBrH118L06 genomic clone, KBrH (HindIII) BAC library Brassica rapa
            subsp. pekinensis, genomic survey sequence.
ACCESSION   CT011342
VERSION     CT011342.1  GI:71466751
KEYWORDS    GSS.
SOURCE      Brassica rapa subsp. pekinensis
  ORGANISM  Brassica rapa subsp. pekinensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1
  AUTHORS   Viehoever,P., Holtgraewe,D. and Weisshaar,B.
  TITLE     BAC end sequences of Brassica rapa
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 917)
  AUTHORS   Li,Y. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-AUG-2005) Weisshaar B., Bielefeld University,
            Institute for Genome Research, Universitaetsstrasse 25, D-33594
            Bielefeld, Germany
COMMENT     Contact: Bernd Weisshaar
            Bielefeld University, Institute for Genome Research
            Universitaetsstrasse 25, D-33594 Bielefeld, Germany Email:
            bernd.weisshaar@uni-bielefeld.de
            BAC end sequences of Brassica rapa BAC clone KBrH118L06; generated
            as contribution to the 'Multinational Brassica rapa Sequencing
            Project' Seq primer: sp6B ATTTAGGTGACACTATAG
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .917
                     /organism="Brassica rapa subsp. pekinensis"
                     /mol_type="genomic DNA"
                     /strain="Chiifu type 401-42"
```

```
                     /cultivar="Chiifu"
                     /sub_species="pekinensis"
                     /db_xref="taxon:51351"
                     /clone="KBrH118L06"
                     /clone_lib="KBrH, Brassica rapa HindIII BAC library
                     GF-SCF-1002, Vector: pCUGIBac1"
                     /lab_host="E.coli DH10B"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 11;  Length 917;
  Best Local Similarity   100.0%;  Pred. No. 55;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 TGGAGCTCAGGCATGAGC 24
              ||||||||||||||||||
Db        524 TGGAGCTCAGGCATGAGC 507


RESULT 13
CC298613
LOCUS       CC298613                 931 bp    DNA     linear   GSS 13-MAY-2003
DEFINITION  CH261-178O6_Sp6.1 CH261 Gallus gallus genomic clone CH261-178O6,
            genomic survey sequence.
ACCESSION   CC298613
VERSION     CC298613.1  GI:30670054
KEYWORDS    GSS.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 931)
  AUTHORS   Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
            Warren,W., Graves,T., Mardis,E. and Wilson,R.
  TITLE     Gallus gallus BAC End Reads
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K. Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 182000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 26
            High quality sequence stop: 790.
FEATURES             Location/Qualifiers
     source          1. .931
                     /organism="Gallus gallus"
                     /mol_type="genomic DNA"
                     /strain="Red Jungle Fowl"
                     /db_xref="taxon:9031"
                     /clone="CH261-178O6"
                     /sex="female"
                     /cell_line="UCD001, inbred 256"
                     /clone_lib="CH261"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CH261 Female Chicken library - for library and clone
                     ordering information: http://www.chori.org/bacpac"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 9;  Length 931;
  Best Local Similarity   100.0%;  Pred. No. 55;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 CTCTGGAGCTCAGGCATG 21
              ||||||||||||||||||
Db        535 CTCTGGAGCTCAGGCATG 552


RESULT 14
CC221775
LOCUS       CC221775                1055 bp    DNA     linear   GSS 12-MAY-2003
DEFINITION  CH261-92I22_RM1.1 CH261 Gallus gallus genomic clone CH261-92I22,
            genomic survey sequence.
ACCESSION   CC221775
VERSION     CC221775.1  GI:30546334
KEYWORDS    GSS.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 1055)
  AUTHORS   Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
            Warren,W., Graves,T., Mardis,E. and Wilson,R.
  TITLE     Gallus gallus BAC End Reads
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K. Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 182000   Std Error: 0.00
            Seq primer: RM1 TACGACTCACTATAGGGAGA
            Class: BAC ends
            High quality sequence start: 38
            High quality sequence stop: 708.
FEATURES             Location/Qualifiers
     source          1. .1055
                     /organism="Gallus gallus"
                     /mol_type="genomic DNA"
                     /strain="Red Jungle Fowl"
                     /db_xref="taxon:9031"
                     /clone="CH261-92I22"
                     /sex="female"
                     /cell_line="UCD001, inbred 256"
                     /clone_lib="CH261"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CH261 Female Chicken library - for library and clone
                     ordering information: http://www.chori.org/bacpac"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 9;  Length 1055;
  Best Local Similarity   100.0%;  Pred. No. 55;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 CTCTGGAGCTCAGGCATG 21
              ||||||||||||||||||
Db        531 CTCTGGAGCTCAGGCATG 548


RESULT 15
BU192345
LOCUS       BU192345                1178 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7930941 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6008244
            5', mRNA sequence.
ACCESSION   BU192345
VERSION     BU192345.1  GI:22706320
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1178)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
```

```
            Plate: LLAM13192  row: m  column: 13
            High quality sequence start: 2
            High quality sequence stop: 111.
FEATURES              Location/Qualifiers
     source           1. .1178
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:6008244"
                      /tissue_type="large cell carcinoma"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="NIH_MGC_68"
                      /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                      Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                      Average insert size 1.8 kb.   Library constructed by Life
                      Technologies. "
ORIGIN

  Query Match             62.1%;  Score 18;  DB 5;  Length 1178;
  Best Local Similarity   100.0%;  Pred. No. 56;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 CTGGAGCTCAGGCATGAG 23
              ||||||||||||||||||
Db        258 CTGGAGCTCAGGCATGAG 275


RESULT 16
DN895382
LOCUS       DN895382                 260 bp    mRNA    linear   EST 25-APR-2005
DEFINITION  nao66b10.y1 Zebrafish Posterior segment. Unnormalized (nao) Danio
            rerio cDNA clone nao66b10 5', mRNA sequence.
ACCESSION   DN895382
VERSION     DN895382.1  GI:62880145
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 260)
  AUTHORS   Wistow,G.
  TITLE     NEIBank analysis of Zebrafish Posterior segment
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 66  row: b  column: 10
            Seq primer: Universal M13 Reverse.
FEATURES              Location/Qualifiers
     source           1. .260
                      /organism="Danio rerio"
                      /mol_type="mRNA"
                      /db_xref="taxon:7955"
                      /clone="nao66b10"
                      /tissue_type="Posterior segment"
                      /dev_stage="Adult"
                      /lab_host="EMDH10B"
                      /clone_lib="Zebrafish Posterior segment. Unnormalized
                      (nao)"
                      /note="Organ: Eye; Vector: pCMVSport6; RNA was extracted
                      from Zebrafish posterior segment tissue (with most retina
                      removed). A directionally cloned cDNA library in the
                      pCMVSPORT6 vector (Invitrogen) was constructed at Bioserve
                      Biotechnology (Laurel MD) essentially following the
                      protocols of the SuperScript Plasmid System, full details
                      of which are contained in the manufacturer's Instruction
                      manual (http://www.lifetech.com/). First strand synthesis
                      was carried out using a Not I primer-adapter
                      [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. cDNA was
                      cloned in Not I/Sal I sites. EST analysis was performed at
                      the NIH Intramural Sequencing Center (NISC). Analyzed data
                      available through http://neibank.nei.nih.gov."
ORIGIN

  Query Match             58.6%;  Score 17;  DB 8;  Length 260;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CTCTCTGGAGCTCAGGC 18
              |||||||||||||||||
Db        168 CTCTCTGGAGCTCAGGC 184


RESULT 17
BB436030/c
LOCUS       BB436030                 300 bp    mRNA    linear   EST 02-AUG-2000
DEFINITION  BB436030 RIKEN full-length enriched, adult pancreas Islet cells Mus
            musculus cDNA clone C820016K20 3', mRNA sequence.
ACCESSION   BB436030
VERSION     BB436030.1  GI:9275757
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 300)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T.,
            Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
            Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
            Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
            Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C.,
            Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H.,
            Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K.,
            Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A.,
            Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A.,
            Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A.,
            Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S.,
            Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
             Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
            Okazaki,Y. and Hayashizaki,Y.
             Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
             Carninci,P. and Hayashizaki,Y.
             High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
             Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES              Location/Qualifiers
     source           1. .300
                      /organism="Mus musculus"
                      /mol_type="mRNA"
                      /db_xref="taxon:10090"
                      /clone="C820016K20"
```

```
                     /tissue_type="pancreas"
                     /cell_type="Islet cells"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RIKEN full-length enriched, adult pancreas
                     Islet cells"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA
                     was cleaved with XhoI and BamHI. Vector: a modified
                     pBluescript KS(+) after bulk excision from Lambda FLC
                     I.-Islet cells were provided by Hiroo Iwata, Institute for
                     Frontier Medical Sciences, Kyoto University, Sakyo-ku,
                     Kyoto, 606-8507 Japan, whose assistance we gratefully
                     acknowledge."
ORIGIN

  Query Match             58.6%;  Score 17;  DB 2;  Length 300;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 TCTGGAGCTCAGGCATG 21
              |||||||||||||||||
Db        142 TCTGGAGCTCAGGCATG 126


RESULT 18
CF170455
LOCUS       CF170455                377 bp    mRNA    linear   EST 25-JUL-2003
DEFINITION  B0827H12-5 NIA Mouse Newborn Kidney cDNA Library (Long 1) Mus
            musculus cDNA clone NIA:B0827H12 IMAGE:30470111 5', mRNA sequence.
ACCESSION   CF170455
VERSION     CF170455.1  GI:33280004
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 377)
  AUTHORS   Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
  TITLE     Construction of long-transcript enriched cDNA libraries from
            submicrogram amounts of total RNAs by a universal PCR amplification
            method
  JOURNAL   Genome Res. 11 (9), 1553-1558 (2001)
   PUBMED   11544199
COMMENT     Contact: Dawood B. Dudekula
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: B0827  row: H  column: 12
            Seq primer: M13 Reverse
            High quality sequence stop: 377
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .377
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:B0827H12-5"
                     /db_xref="taxon:10090"
                     /clone="NIA:B0827H12 IMAGE:30470111"
                     /dev_stage="Newborn Kidney"
                     /lab_host="DH10B"
                     /clone_lib="NIA Mouse Newborn Kidney cDNA Library (Long
                     1)"
                     /note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
                     Site_2: NotI; Mouse cDNA project by the Laboratory of
                     Genetics, National Institute on Aging (NIA), Intramural
                     Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
                     In brief, double-stranded cDNAs were synthesized with an
                     Oligo(dT) primer [Invitrogen:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT-3'] from
                     26 ug of total RNA, treated with T4 DNA polymerase, and
                     purified by ethanol-precipitation. The cDNAs were ligated
                     to Lone-linker LL-Sal4, purified by phenol/chloroform, and
                     separated from free linkers by Centricon 100. Then, the
                     cDNAs were amplified by long-range high fidelity PCR using
                     Ex Taq polymerase (Takara) with a primer Sal4-S. The
                     products were purified by phenol/chloroform and Centricon
                     100. The cDNAs were digested with SalI and NotI enzymes
                     and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
                     vector. The DH10B E. coli host was transformed with the
                     ligation mixture by the standard chemical method. The
                     average insert size is about 3.0 kb. The library was
                     constructed by Yulan Piao."
ORIGIN

  Query Match             58.6%;  Score 17;  DB 6;  Length 377;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db        222 TCAGGCATGAGCCAGCA 238


RESULT 19
BE753088
LOCUS       BE753088                552 bp    mRNA    linear   EST 25-APR-2001
DEFINITION  205980 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BE753088
VERSION     BE753088.1  GI:10167080
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 552)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
            Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
            Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F.,
            Quackenbush,J. and Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
   PUBMED   11282978
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 47  row: C  column: 8
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .552
```

```
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /clone_lib="MARC 2BOV"
                     /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                     Library made from pooled tissue from testis, thymus,
                     semitendonosus muscle, longissimus muscle, pancreas,
                     adrenal, and endometrium."
ORIGIN

  Query Match            58.6%;  Score 17;  DB 2;  Length 552;
  Best Local Similarity  100.0%;  Pred. No. 1.8e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAGG 17
              |||||||||||||||||
Db         22 CCTCTCTGGAGCTCAGG 38


RESULT 20
BI889195
LOCUS       BI889195                611 bp    mRNA    linear   EST 12-OCT-2001
DEFINITION  ZF637-2-000950 Zebrafish shield stage whole embryo cDNA library
            MPMGp637 Danio rerio cDNA clone MPMGp637_22E18;MPMGp637E1822 5',
            mRNA sequence.
ACCESSION   BI889195
VERSION     BI889195.1  GI:16096466
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 611)
  AUTHORS   Clark,M., Aanstad,P., Hennig,S., Johnson,S.L. and Lehrach,H.
  TITLE     EST sequencing of a zebrafish shield stage cDNA library normalised
            by oligonucleotide fingerprinting
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Hennig S
            laboraty 123, dept.Lehrach
            Max-Planck-Institut fuer Molekulare Genetik
            Ihnestr.63-73, D-14195 Berlin, Germany
            Tel: +49 30 8413 1612
            Fax: +49 30 8413 1380
            Email: hennig@molgen.mpg.de
            5' EST sequencing of clones from a zebrafish shield stage library,
            normalised from 55,000 starting clones by oligonucleotide
            fingerprinting
            High quality sequence stop: 611.
FEATURES             Location/Qualifiers
     source          1. .611
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="MPMGp637_22E18;MPMGp637E1822"
                     /tissue_type="whole embryo"
                     /dev_stage="shield stage, 6 hrs post-fertilisation"
                     /lab_host="E.coli, XL1 blue MRF"
                     /clone_lib="Zebrafish shield stage whole embryo cDNA
                     library MPMGp637"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI;
                     oligo-dT-NotI primed, SalI adaptors, directionally cloned,
                     library normalised by oligonucleotide fingerprinting"
ORIGIN

  Query Match            58.6%;  Score 17;  DB 3;  Length 611;
  Best Local Similarity  100.0%;  Pred. No. 1.9e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CTCTCTGGAGCTCAGGC 18
              |||||||||||||||||

Db        534 CTCTCTGGAGCTCAGGC 550


RESULT 21
CF169566
LOCUS       CF169566                676 bp    mRNA    linear   EST 25-JUL-2003
DEFINITION  B0815D08-5 NIA Mouse Newborn Kidney cDNA Library (Long 1) Mus
            musculus cDNA clone NIA:B0815D08 IMAGE:30468907 5', mRNA sequence.
ACCESSION   CF169566
VERSION     CF169566.1  GI:33279115
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 676)
  AUTHORS   Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
  TITLE     Construction of long-transcript enriched cDNA libraries from
            submicrogram amounts of total RNAs by a universal PCR amplification
            method
  JOURNAL   Genome Res. 11 (9), 1553-1558 (2001)
   PUBMED   11544199
COMMENT     Contact: Dawood B. Dudekula
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: B0815  row: D  column: 08
            Seq primer: M13 Reverse
            High quality sequence stop: 676
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .676
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:B0815D08-5"
                     /db_xref="taxon:10090"
                     /clone="NIA:B0815D08 IMAGE:30468907"
                     /dev_stage="Newborn Kidney"
                     /lab_host="DH10B"
                     /clone_lib="NIA Mouse Newborn Kidney cDNA Library (Long
                     1)"
                     /note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
                     Site_2: NotI; Mouse cDNA project by the Laboratory of
                     Genetics, National Institute on Aging (NIA), Intramural
                     Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
                     In brief, double-stranded cDNAs were synthesized with an
                     Oligo(dT) primer [Invitrogen:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT-3'] from
                     26 ug of total RNA, treated with T4 DNA polymerase, and
                     purified by ethanol-precipitation. The cDNAs were ligated
                     to Lone-linker LL-Sal4, purified by phenol/chloroform, and
                     separated from free linkers by Centricon 100. Then, the
                     cDNAs were amplified by long-range high fidelity PCR using
                     Ex Taq polymerase (Takara) with a primer Sal4-S. The
                     products were purified by phenol/chloroform and Centricon
                     100. The cDNAs were digested with SalI and NotI enzymes
                     and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
                     vector. The DH10B E. coli host was transformed with the
                     ligation mixture by the standard chemical method. The
                     average insert size is about 3.0 kb. The library was
                     constructed by Yulan Piao."
ORIGIN

  Query Match            58.6%;  Score 17;  DB 6;  Length 676;
  Best Local Similarity  100.0%;  Pred. No. 1.9e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db        367 TCAGGCATGAGCCAGCA 383
```

```
RESULT 22
CV224516/c
LOCUS       CV224516                718 bp    mRNA    linear   EST 21-SEP-2004
DEFINITION  CS_hyp_16c09_M13Reverse Blue crab hypodermis, normalized
            Callinectes sapidus cDNA clone CS_hyp_16c09 5' similar to
            ref|NP_001001478.1| complement related-long precursor -
            Strongylocentrotus purpuratus. Score = 33.1 bits (74), Expect =
            5.7, mRNA sequence.
ACCESSION   CV224516
VERSION     CV224516.1  GI:52371767
KEYWORDS    EST.
SOURCE      Callinectes sapidus (blue crab)
  ORGANISM  Callinectes sapidus
            Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
            Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
            Eubrachyura; Portunoidea; Portunidae; Callinectes.
REFERENCE   1  (bases 1 to 718)
  AUTHORS   Shafer,T.H., Coblentz,F.E. and Towle,D.W.
  TITLE     Expressed sequence tags from normalized cDNA libraries prepared
            from gill and hypodermis tissues of the blue crab, Callinectes
            sapidus
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Thomas H. Shafer
            Department of Biological Sciences
            University of North Carolina Wilmington
            601 S. College Rd, Wilmington, NC 28403, USA
            Tel: 910-962-7275
            Fax: 910-962-4066
            Email: shafert@uncw.edu
            Plate: 16  row: c  column: 09
            Seq primer: M13 Reverse
            High quality sequence stop: 491.
FEATURES             Location/Qualifiers
     source          1. .718
                     /organism="Callinectes sapidus"
                     /mol_type="mRNA"
                     /db_xref="taxon:6763"
                     /clone="CS_hyp_16c09"
                     /tissue_type="Pooled hypodermal epithelium from the
                     mid-dorsal region and arthrodial membrane of premolt
                     (stage D2) and 3-hour postmolt crabs"
                     /dev_stage="Adult"
                     /clone_lib="Blue crab hypodermis, normalized"
                     /note="Vector: pCMV Sport 6.1; Total RNA samples were
                     prepared individually from each tissue, checked for
                     quality, and then pooled for construction and
                     normalization of a cDNA library by Invitrogen. Plasmids
                     were isolated and inserts sequenced from their 5'-ends by
                     the Blue Crab Molecular Genetics Laboratory at the
                     University of North Carolina Wilmington. Traces were
                     trimmed, compared (BLASTx) to NCBI non-redundant protein
                     database as of 19 July 2004, and processed for submission
                     to dbEST by trace2dbEST software (Parkinson, Anthony and
                     Blaxter, unpublished software)."
ORIGIN

  Query Match             58.6%;  Score 17;  DB 7;  Length 718;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 CTGGAGCTCAGGCATGA 22
              |||||||||||||||||
Db        209 CTGGAGCTCAGGCATGA 193


RESULT 23
AG490318/c
LOCUS       AG490318                802 bp    DNA     linear   GSS 22-DEC-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-387K02.T7, genomic survey
            sequence.
ACCESSION   AG490318
VERSION     AG490318.1  GI:48197548
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus (Japanese wild mouse)
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Abe,K., Noguchi,H., Tagawa,K., Yuzuriha,M., Toyoda,A., Kojima,T.,
            Ezawa,K., Saitou,N., Hattori,M., Sakaki,Y., Moriwaki,K. and
            Shiroishi,T.
  TITLE     Contribution of Asian mouse subspecies Mus musculus molossinus to
            genomic constitution of strain C57BL/6J, as defined by BAC-end
            sequence-SNP analysis
  JOURNAL   Genome Res. 14 (12), 2439-2447 (2004)
   PUBMED   15574823
REFERENCE   2  (bases 1 to 802)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institude, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : T7
            LIBRARY
            Vector     : pBACe3.6
            R.Site 1   : EcoRI
            R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .802
                     /organism="Mus musculus molossinus"
                     /mol_type="genomic DNA"
                     /sub_species="molossinus"
                     /db_xref="taxon:57486"
                     /clone="MSMg01-387K02.T7"
                     /sex="male"
                     /tissue_type="mixture of kidney and spleen"
                     /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 10;  Length 802;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db        312 TCAGGCATGAGCCAGCA 296


RESULT 24
AQ894819
LOCUS       AQ894819                809 bp    DNA     linear   GSS 10-NOV-1999
DEFINITION  HS_3133_A1_E07_T7C CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3133 Col=13 Row=I, genomic survey
            sequence.
ACCESSION   AQ894819
VERSION     AQ894819.1  GI:6351009
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 809)
```

```
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 3133  row: I  column: 13
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 809.
FEATURES             Location/Qualifiers
     source          1. .809
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3133 Col=13 Row=I"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 9;  Length 809;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db        648 TCAGGCATGAGCCAGCA 664


RESULT 25
AZ752023
LOCUS       AZ752023                818 bp    DNA     linear   GSS 25-JAN-2001
DEFINITION  RPCI-24-115N13.TV RPCI-24 Mus musculus genomic clone
            RPCI-24-115N13, genomic survey sequence.
ACCESSION   AZ752023
VERSION     AZ752023.1  GI:12537182
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 818)
  AUTHORS   Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
            Russell,D., de Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-24
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-24. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 115  row: N  column: 13
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .818
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-24-115N13"
                     /sex="Male"
                     /cell_type="Spleen/Brain"
                     /clone_lib="RPCI-24"
                     /note="Vector: pTARBAC1; Site_1: BamH1; Site_2: BamH1;
                     RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
                     library was  cloned in the pTARBAC1 cloning vector at the
                     BamH1 sites using MboI  partially digested male C57BL/6J
                     DNA."
ORIGIN

  Query Match             58.6%;  Score 17;  DB 9;  Length 818;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAGG 17
              |||||||||||||||||
Db        413 CCTCTCTGGAGCTCAGG 429


RESULT 26
CC475574/c
LOCUS       CC475574                918 bp    DNA     linear   GSS 16-JUN-2003
DEFINITION  CH240_301I3.T7 CHORI-240 Bos taurus genomic clone CH240_301I3,
            genomic survey sequence.
ACCESSION   CC475574
VERSION     CC475574.1  GI:31752691
KEYWORDS    GSS.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 918)
  AUTHORS   Holt,R., Stott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,
            Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,
            Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
            Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
            Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
            Dalrymple,B.P. and Tellam,R.
  TITLE     Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: CH240_301I3.TARBAC13P2
            Contact: Rob Holt
            Sequencing
            The British Columbia Cancer Agency Genome Science Centre
            600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
            Tel: 604-877-6085
            Fax: 604-877-6276
            Email: rholt@bcgsc.ca
            Clones are derived from the bovine BAC library CHORI-240
            (http://www.chori.org/bacpac/bovine240.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/ordering_information.htm). This work
            was undertaken as part of the International Bovine BAC Mapping
            Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
            British Columbia Genome Sciences Centre, Canada.
            Plate: 301  row: I  column: 3
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .918
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
```

```
                     /strain="breed: Hereford"
                     /db_xref="taxon:9913"
                     /clone="CH240_301I3"
                     /sex="Male"
                     /cell_type="Blood"
                     /clone_lib="CHORI-240"
                     /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
                     Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
                     library (Male) produced by Pieter de Jong"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 9;  Length 918;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAGG 17
              |||||||||||||||||
Db        146 CCTCTCTGGAGCTCAGG 130


RESULT 27
CC189709
LOCUS       CC189709                1104 bp    DNA     linear   GSS 08-MAY-2003
DEFINITION  CH261-131C19_RM1.1 CH261 Gallus gallus genomic clone CH261-131C19,
            genomic survey sequence.
ACCESSION   CC189709
VERSION     CC189709.1  GI:30434222
KEYWORDS    GSS.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 1104)
  AUTHORS   Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
            Warren,W., Graves,T., Mardis,E. and Wilson,R.
  TITLE     Gallus gallus BAC End Reads
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K. Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 182000   Std Error: 0.00
            Seq primer: RM1 TACGACTCACTATAGGGAGA
            Class: BAC ends
            High quality sequence start: 11
            High quality sequence stop: 330.
FEATURES             Location/Qualifiers
     source          1. .1104
                     /organism="Gallus gallus"
                     /mol_type="genomic DNA"
                     /strain="Red Jungle Fowl"
                     /db_xref="taxon:9031"
                     /clone="CH261-131C19"
                     /sex="female"
                     /cell_line="UCD001, inbred 256"
                     /clone_lib="CH261"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CH261 Female Chicken library - for library and clone
                     ordering information: http://www.chori.org/bacpac"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 9;  Length 1104;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 TGGAGCTCAGGCATGAG 23
              |||||||||||||||||
Db        398 TGGAGCTCAGGCATGAG 414


RESULT 28
CR261805
LOCUS       CR261805                  70 bp    DNA     linear   GSS 06-JUL-2004
DEFINITION  Forward strand read from insert in 5'HPRT insertion targeting and
            chromosome engineering clone MHPN366p08, genomic survey sequence.
ACCESSION   CR261805
VERSION     CR261805.1  GI:50040658
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPN366p08"
                     /clone_lib="MHPN"
ORIGIN

  Query Match             55.2%;  Score 16;  DB 11;  Length 70;
  Best Local Similarity   100.0%;  Pred. No. 5.7e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CTCTCTGGAGCTCAGG 17
              ||||||||||||||||
Db         16 CTCTCTGGAGCTCAGG 31


RESULT 29
CR058210
LOCUS       CR058210                  77 bp    DNA     linear   GSS 05-JUL-2004
DEFINITION  Forward strand read from insert in 5'HPRT insertion targeting and
            chromosome engineering clone MHPN344n05, genomic survey sequence.
ACCESSION   CR058210
VERSION     CR058210.1  GI:49791682
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 77)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .77
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPN344n05"
                     /clone_lib="MHPN"
ORIGIN

  Query Match             55.2%;  Score 16;  DB 11;  Length 77;
  Best Local Similarity   100.0%;  Pred. No. 5.7e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CTCTCTGGAGCTCAGG 17
              ||||||||||||||||
Db          9 CTCTCTGGAGCTCAGG 24
```

```
RESULT 30
N88936
LOCUS       N88936                 120 bp    mRNA    linear   EST 02-APR-1996
DEFINITION  K6757F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA
            clone K6757 5' similar to REPETITIVE ELEMENT ALU, mRNA sequence.
ACCESSION   N88936
VERSION     N88936.1  GI:1442266
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 120)
  AUTHORS   Liew,C.C.
  TITLE     cDNAs from fetal heart (1996)
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Liew CC
            Brigham and Women's Hospital
            Harvard Medical School
            75 Francis St. Boston, MA 02115, USA
            Tel: 6177328915
            Fax: 6179750995
            Email: cliew@rics.bwh.harvard.edu
            Seq primer: GAAATTAACCCTCACTAAAGGG.
FEATURES             Location/Qualifiers
     source          1. .120
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="K6757"
                     /lab_host="E. coli XL1-Blue"
                     /clone_lib="Human fetal heart, Lambda ZAP Express"
                     /note="Vector: Lambda ZAP Express; Site_1: EcoRI; Site_2:
                     XhoI; mRNA was purified from human fetal hearts (8-10
                     weeks). cDNA was synthesized using a XhoI-Oligo dT
                     adaptor-primer. EcoRI adaptors were ligated, followed by
                     digestion with XhoI, for directional cloning into
                     predigested lambda ZAP Express."
ORIGIN

  Query Match           55.2%;  Score 16;  DB 8;  Length 120;
  Best Local Similarity  100.0%;  Pred. No. 5.9e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db         18 CAGGCATGAGCCAGCA 33


RESULT 31
BF927034/c
LOCUS       BF927034               143 bp    mRNA    linear   EST 19-JAN-2001
DEFINITION  CM2-NT0192-051200-577-a11 NT0192 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF927034
VERSION     BF927034.1  GI:12324918
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 143)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM2&t2=CM2-NT0192-
            051200-577-a11&t3=2000-12-05&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 21
            High quality sequence stop: 143.
FEATURES             Location/Qualifiers
     source          1. .143
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="NT0192"
                     /note="Organ: nervous_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match           55.2%;  Score 16;  DB 2;  Length 143;
  Best Local Similarity  100.0%;  Pred. No. 5.9e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db         40 CAGGCATGAGCCAGCA 25


RESULT 32
BF800261
LOCUS       BF800261               150 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  CM4-CI0061-181000-368-b11 CI0061 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF800261
VERSION     BF800261.1  GI:12129250
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 150)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
```

```
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM4&t2=CM4-CI0061-
            181000-368-b11&t3=2000-10-18&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 20
            High quality sequence stop: 147.
FEATURES             Location/Qualifiers
     source          1. .150
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="CI0061"
                     /note="Organ: colon_ins; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             55.2%;  Score 16;  DB 2;  Length 150;
  Best Local Similarity   100.0%;  Pred. No. 5.9e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db         36 CAGGCATGAGCCAGCA 51


RESULT 33
BF882488/c
LOCUS       BF882488                168 bp    mRNA    linear   EST 17-JAN-2001
DEFINITION  CM1-ET0191-051200-626-b01 ET0191 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF882488
VERSION     BF882488.1  GI:12272614
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 168)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM1&t2=CM1-ET0191-
            051200-626-b01&t3=2000-12-05&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 167.
FEATURES             Location/Qualifiers
     source          1. .168
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="ET0191"
                     /note="Organ: lung_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             55.2%;  Score 16;  DB 2;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 6e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CTCTCTGGAGCTCAGG 17
              ||||||||||||||||
Db        124 CTCTCTGGAGCTCAGG 109


RESULT 34
BB720345
LOCUS       BB720345                181 bp    mRNA    linear   EST 12-OCT-2001
DEFINITION  BB720345 RIKEN full-length enriched, adult male liver tumor Mus
            musculus cDNA clone C730034M09 3', mRNA sequence.
ACCESSION   BB720345
VERSION     BB720345.1  GI:16101918
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 181)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
            Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
            Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
            Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
            and Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
            Sugahara,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
```

```
          e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .181
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="C730034M09"
                     /sex="male"
                     /tissue_type="liver tumor"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RIKEN full-length enriched, adult male liver
                     tumor"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA
                     was cleaved with BamHI and XhoI. Vector: a modified
                     pBluescript KS(+) after bulk excision from Lambda FLC I.
                     Tissue was provided by William A. Held, Roswell Park
                     Cancer Institute, Department of Molecular and Cellular
                     Biology, Elm and Carlton Streets, Buffalo, NY 14263, whose
                     assistance we gratefully acknowledge."
ORIGIN

  Query Match             55.2%;  Score 16;  DB 2;  Length 181;
  Best Local Similarity   100.0%;  Pred. No. 6e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         12 CTCAGGCATGAGCCAG 27
              ||||||||||||||||
Db         49 CTCAGGCATGAGCCAG 64


RESULT 35
CE664705/c
LOCUS       CE664705                191 bp    DNA     linear   GSS 29-SEP-2003
DEFINITION  tigr-gss-dog-17000329162077 Dog Library Canis familiaris genomic,
            genomic survey sequence.
ACCESSION   CE664705
VERSION     CE664705.1  GI:36983573
KEYWORDS    GSS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
            Canis.
REFERENCE   1  (bases 1 to 191)
  AUTHORS   Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
  TITLE     The dog genome: survey sequencing and comparative analysis
  JOURNAL   Science 301 (5641), 1898-1903 (2003)
   PUBMED   14512627
COMMENT     Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirknes@tigr.org
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .191
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Standard Poodle"
                     /db_xref="taxon:9615"
                     /clone_lib="Dog Library"
                     /note="Site_1: BstXI; Libraries were prepared from
                     peripheral blood"
ORIGIN

  Query Match             55.2%;  Score 16;  DB 10;  Length 191;
  Best Local Similarity   100.0%;  Pred. No. 6e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAG 16
              ||||||||||||||||
Db        150 CCTCTCTGGAGCTCAG 135


RESULT 36
B82265/c
LOCUS       B82265                  222 bp    DNA     linear   GSS 09-APR-1999
DEFINITION  RPCI11-14D21.TP RPCI-11 Homo sapiens genomic clone RPCI-11-14D21,
            genomic survey sequence.
ACCESSION   B82265
VERSION     B82265.1  GI:2869288
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 222)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences for Sequence-Ready Map Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .222
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7505084"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-14D21"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             55.2%;  Score 16;  DB 9;  Length 222;
  Best Local Similarity   100.0%;  Pred. No. 6e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db         73 CAGGCATGAGCCAGCA 58
```

```
RESULT 37
AZ331080/c
LOCUS       AZ331080                239 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0056B17R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0056B17 R, genomic survey sequence.
ACCESSION   AZ331080
VERSION     AZ331080.1  GI:10393258
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 239)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0056  row: B  column: 17
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 239.
FEATURES             Location/Qualifiers
     source          1. .239
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0056B17"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             55.2%;  Score 16;  DB 9;  Length 239;
  Best Local Similarity   100.0%;  Pred. No. 6.1e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CTCTCTGGAGCTCAGG 17
              ||||||||||||||||
Db        143 CTCTCTGGAGCTCAGG 128


RESULT 38
AZ800758/c
LOCUS       AZ800758                252 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0058N22R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0058N22 R, genomic survey sequence.
ACCESSION   AZ800758
VERSION     AZ800758.1  GI:12953081
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 252)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0058  row: N  column: 22
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 252.
FEATURES             Location/Qualifiers
     source          1. .252
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0058N22"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             55.2%;  Score 16;  DB 9;  Length 252;
  Best Local Similarity   100.0%;  Pred. No. 6.1e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
```

```
                ||||||||||||||||
Db        151 CAGGCATGAGCCAGCA 136


RESULT 39
CV326948
LOCUS       CV326948                 266 bp    mRNA    linear   EST 24-SEP-2004
DEFINITION  CM4-UT0042-080900-307-a09 UT0042 Homo sapiens cDNA, mRNA sequence.
ACCESSION   CV326948
VERSION     CV326948.1  GI:52650162
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 266)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. http://www.ludwig.org.br.
FEATURES             Location/Qualifiers
     source          1. .266
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="UT0042"
                     /note="Organ: uterus_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             55.2%;  Score 16;  DB 7;  Length 266;
  Best Local Similarity   100.0%;  Pred. No. 6.1e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db         84 CAGGCATGAGCCAGCA 99


RESULT 40
BI024787/c
LOCUS       BI024787                 267 bp    mRNA    linear   EST 14-JUN-2001
DEFINITION  PM0-MT0202-300101-002-a11 MT0202 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI024787
VERSION     BI024787.1  GI:14431417
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 267)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM0&t2=PM0-MT0202-
            300101-002-a11&t3=2001-01-30&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 3
            High quality sequence stop: 267.
FEATURES             Location/Qualifiers
     source          1. .267
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="MT0202"
                     /note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No.
                     196,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
ORIGIN

  Query Match             55.2%;  Score 16;  DB 2;  Length 267;
  Best Local Similarity   100.0%;  Pred. No. 6.1e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        256 CAGGCATGAGCCAGCA 241


RESULT 41
AA579864
LOCUS       AA579864                 270 bp    mRNA    linear   EST 03-SEP-1997
DEFINITION  nj41b10.s1 NCI_CGAP_AA1 Homo sapiens cDNA clone IMAGE:995035 3'
            similar to contains Alu repetitive element;contains L1.t3 L1
            repetitive element ;, mRNA sequence.
ACCESSION   AA579864
VERSION     AA579864.1  GI:2355191
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 270)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
```

```
            Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
            Ph.D.
             cDNA Library Arraying: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 264.
FEATURES              Location/Qualifiers
     source           1. .270
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:995035"
                      /tissue_type="adrenal adenoma"
                      /lab_host="SOLR (kanamycin resistant)"
                      /clone_lib="NCI_CGAP_AA1"
                      /note="Organ: adrenal gland; Vector: Bluescript SK-;
                      Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally.
                      Primer: Oligo dT. Two pooled bulk adrenal adenomas. 5'
                      adaptor sequence: 5' GAATTCGGCACGAG 3' 3' adaptor
                      sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3' Average insert
                      size:  1.6 kb."
ORIGIN

  Query Match             55.2%;  Score 16;  DB 1;  Length 270;
  Best Local Similarity   100.0%;  Pred. No. 6.1e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db         97 CAGGCATGAGCCAGCA 112


RESULT 42
AZ065220/c
LOCUS       AZ065220                273 bp    DNA     linear   GSS 30-MAR-2000
DEFINITION  RPCI-23-391I2.TV RPCI-23 Mus musculus genomic clone RPCI-23-391I2,
            genomic survey sequence.
ACCESSION   AZ065220
VERSION     AZ065220.1  GI:7356472
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 273)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S.,
            Akinret,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de
            Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-391I2.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 391  row: I  column: 2
            Seq primer: T7
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .273
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="RPCI-23-391I2"
                      /sex="Female"
                      /lab_host="DH10B"
                      /clone_lib="RPCI-23"
                      /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                      EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                      brain genomic DNA was isolated and  partially digested
                      with a combination of EcoRI and EcoRI Methylase.   Size
                      selected DNA was cloned into the pBACe3.6 vector at the
                      EcoRI  sites.  The ligation products were transformed into
                      DH10B  electrocompetent cells (BRL Life Technologies).   "
ORIGIN

  Query Match             55.2%;  Score 16;  DB 9;  Length 273;
  Best Local Similarity   100.0%;  Pred. No. 6.1e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAG 16
              ||||||||||||||||
Db        209 CCTCTCTGGAGCTCAG 194


RESULT 43
AQ424457
LOCUS       AQ424457                282 bp    DNA     linear   GSS 24-MAR-1999
DEFINITION  CITBI-E1-255701.TR CITBI-E1 Homo sapiens genomic clone 255701,
            genomic survey sequence.
ACCESSION   AQ424457
VERSION     AQ424457.1  GI:4497723
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 282)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .282
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone="255701"
                      /sex="male"
                      /cell_type="sperm"
                      /clone_lib="CITBI-E1"
                      /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                      CalTech Human BAC Library D"
ORIGIN

  Query Match             55.2%;  Score 16;  DB 9;  Length 282;
```

```
  Best Local Similarity   100.0%;  Pred. No. 6.1e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        228 CAGGCATGAGCCAGCA 243


RESULT 44
AQ052047/c
LOCUS       AQ052047                290 bp    DNA     linear   GSS 20-APR-1999
DEFINITION  RPCI11-53B18.TK RPCI-11 Homo sapiens genomic clone RPCI-11-53B18,
            genomic survey sequence.
ACCESSION   AQ052047
VERSION     AQ052047.1  GI:3349084
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 290)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences for Sequence-Ready Map Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: RPCI11-53B18.TJ
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .290
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7520009"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-53B18"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             55.2%;  Score 16;  DB 9;  Length 290;
  Best Local Similarity   100.0%;  Pred. No. 6.1e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        257 CAGGCATGAGCCAGCA 242


RESULT 45
B79172/c
LOCUS       B79172                  298 bp    DNA     linear   GSS 18-JUN-1998
DEFINITION  CIT978SK-11E14.TV CIT978SK Homo sapiens genomic clone 11E14,
            genomic survey sequence.
ACCESSION   B79172
VERSION     B79172.1  GI:2866195
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 298)
  AUTHORS   Kim,U.-J., Adams,M.D. and Simon,M.I.
  TITLE     Determination of clone end sequences of human Bacterial Artificial
            Chromosomes
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT978SK-11E14.TP CIT978SK-11E14.TV.1 CIT-HSP-11E14.TVB
            Contact: Ung-Jin Kim
            CalTech Genome Research Lab
            California Institute of Technology
            Division of Biology, MS 147-75, Pasadena, CA 91125, USA
            Tel: 626 796 7066
            Fax: 626 395 4901
            Email: ung@ash.tree.caltech.edu
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .298
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:5222475"
                     /db_xref="taxon:9606"
                     /clone="11E14"
                     /sex="Female"
                     /cell_type="Fibroblast"
                     /clone_lib="CIT978SK"
                     /note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
                     CalTech Human BAC Library A"
ORIGIN

  Query Match             55.2%;  Score 16;  DB 9;  Length 298;
  Best Local Similarity   100.0%;  Pred. No. 6.1e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        205 CAGGCATGAGCCAGCA 190


RESULT 46
B79288/c
LOCUS       B79288                  298 bp    DNA     linear   GSS 18-JUN-1998
DEFINITION  CIT978SK-11E14.TV.1 CIT978SK Homo sapiens genomic clone 11E14,
            genomic survey sequence.
ACCESSION   B79288
VERSION     B79288.1  GI:2866311
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 298)
  AUTHORS   Kim,U.-J., Adams,M.D. and Simon,M.I.
  TITLE     Determination of clone end sequences of human Bacterial Artificial
            Chromosomes
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT978SK-11E14.TP CIT978SK-11E14.TV CIT-HSP-11E14.TVB
            Contact: Ung-Jin Kim
            CalTech Genome Research Lab
            California Institute of Technology
            Division of Biology, MS 147-75, Pasadena, CA 91125, USA
            Tel: 626 796 7066
            Fax: 626 395 4901
            Email: ung@ash.tree.caltech.edu
            Clones are available from Research Genetics (info@resgen.com). BAC
```

```
          end search page:
          http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
          Seq primer: T7
          Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .298
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:5222475"
                     /db_xref="taxon:9606"
                     /clone="11E14"
                     /sex="Female"
                     /cell_type="Fibroblast"
                     /clone_lib="CIT978SK"
                     /note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
                     CalTech Human BAC Library A"
ORIGIN

  Query Match             55.2%;  Score 16;  DB 9;  Length 298;
  Best Local Similarity   100.0%;  Pred. No. 6.1e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db         205 CAGGCATGAGCCAGCA 190


RESULT 47
BB400771
LOCUS       BB400771                309 bp    mRNA    linear   EST 15-JUL-2000
DEFINITION  BB400771 RIKEN full-length enriched, ES cells Mus musculus cDNA
            clone C330019P13 3' similar to AF213389 Mus musculus ATP-binding
            cassette protein (Abcc1b) mRNA, mRNA sequence.
ACCESSION   BB400771
VERSION     BB400771.1  GI:9220167
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 309)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T.,
            Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
            Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
            Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
            Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C.,
            Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H.,
            Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K.,
            Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A.,
            Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A.,
            Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A.,
            Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S.,
            Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
             Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
            Okazaki,Y. and Hayashizaki,Y.
             Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
             Carninci,P. and Hayashizaki,Y.
             High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
             Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .309
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="C330019P13"
                     /cell_type="ES cells"
                     /lab_host="SOLR"
                     /clone_lib="RIKEN full-length enriched, ES cells"
                     /note="Site_1: XhoI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer (5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'), cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 5.0 and subtraction to Rot = 25.0. Second strand
                     cDNA was prepared with the primer adapter of sequence (5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3')."
ORIGIN

  Query Match             55.2%;  Score 16;  DB 2;  Length 309;
  Best Local Similarity   100.0%;  Pred. No. 6.1e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db         245 CAGGCATGAGCCAGCA 260


RESULT 48
CR069381
LOCUS       CR069381                311 bp    DNA     linear   GSS 05-JUL-2004
DEFINITION  Forward strand read from insert in 5'HPRT insertion targeting and
            chromosome engineering clone MHPN222j23, genomic survey sequence.
ACCESSION   CR069381
VERSION     CR069381.1  GI:49802971
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 311)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .311
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPN222j23"
                     /clone_lib="MHPN"
ORIGIN

  Query Match             55.2%;  Score 16;  DB 11;  Length 311;
  Best Local Similarity   100.0%;  Pred. No. 6.2e+02;
```

```
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CTCTCTGGAGCTCAGG 17
              ||||||||||||||||
Db         42 CTCTCTGGAGCTCAGG 57


RESULT 49
BG546523/c
LOCUS       BG546523                319 bp    mRNA    linear   EST 04-APR-2001
DEFINITION  602574272F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4702696 5',
            mRNA sequence.
ACCESSION   BG546523
VERSION     BG546523.1  GI:13545188
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 319)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1539  row: o  column: 17
            High quality sequence stop: 318.
FEATURES             Location/Qualifiers
     source          1. .319
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4702696"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_77"
                     /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size 1.9
                     kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech  Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             55.2%;  Score 16;  DB 2;  Length 319;
  Best Local Similarity   100.0%;  Pred. No. 6.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db         21 CAGGCATGAGCCAGCA 6


RESULT 50
AI445839
LOCUS       AI445839                320 bp    mRNA    linear   EST 13-APR-1999
DEFINITION  tj12d08.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2141295 3'
            similar to contains Alu repetitive element;contains element PTR5
            repetitive element ;, mRNA sequence.
ACCESSION   AI445839
VERSION     AI445839.1  GI:4291016
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 320)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 2554    Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 318.
FEATURES             Location/Qualifiers
     source          1. .320
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2141295"
                     /tissue_type="poorly differentiated adenocarcinoma with
                     signet ring cell features"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Gas4"
                     /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.69 kb. Life Technologies catalog #:
                     11549-011"
ORIGIN

  Query Match             55.2%;  Score 16;  DB 1;  Length 320;
  Best Local Similarity   100.0%;  Pred. No. 6.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGC 28
              ||||||||||||||||
Db        263 TCAGGCATGAGCCAGC 278


Search completed: April 12, 2006, 14:48:09
Job time : 3767 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: April 12, 2006, 12:55:23 ; Search time 311 Seconds
(without alignments)
621.467 Million cell updates/sec

Title: SEQ1-4023-4051-4037A
Perfect score: 29
Sequence: 1 cctctctggagctcaggcatgagccagca 29

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4996997 seqs, 3332346308 residues

Word size : 15

Total number of hits satisfying chosen parameters: 329

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	19	65.5	2145	4	AAS15366	Aas15366 cDNA enco
c	2	17	58.6	21728	12	ADQ97385	Adq97385 Human can
c	3	16	55.2	391	5	ABV56351	Abv56351 Human pro
	4	16	55.2	394	9	ACH17384	Ach17384 Human adu
	5	16	55.2	410	5	ABV58668	Abv58668 Human pro
	6	16	55.2	426	5	ABV18799	Abv18799 Human pro
c	7	16	55.2	429	9	ACH14691	Ach14691 Human adu
c	8	16	55.2	443	9	ACH25630	Ach25630 Human adu
c	9	16	55.2	450	9	ACH24672	Ach24672 Human adu
c	10	16	55.2	452	6	ABN18360	Abn18360 Human ORF
	11	16	55.2	463	4	AAI12911	Aai12911 Probe #28
	12	16	55.2	463	4	ABA54612	Aba54612 Human foe
	13	16	55.2	463	4	AAI34271	Aai34271 Probe #29
	14	16	55.2	463	4	ABA44163	Aba44163 Human bre
	15	16	55.2	463	4	ABA24397	Aba24397 Probe #28
	16	16	55.2	463	4	AAK28345	Aak28345 Human bon
	17	16	55.2	463	4	AAK02903	Aak02903 Human bra
	18	16	55.2	463	4	ABS27944	Abs27944 Human liv
	19	16	55.2	463	5	AAI02830	Aai02830 Probe #28
	20	16	55.2	463	6	ABS02854	Abs02854 Human gen
c	21	16	55.2	495	5	ABV50399	Abv50399 Human pro
	22	16	55.2	564	5	ABV48581	Abv48581 Human pro
c	23	16	55.2	573	6	ABN63733	Abn63733 Human can
	24	16	55.2	601	14	AEB33143	Aeb33143 Human DNA
c	25	16	55.2	768	9	ADB81754	Adb81754 Human cDN
	26	16	55.2	809	4	AAH03489	Aah03489 Human cDN
c	27	16	55.2	817	6	ABQ88810	Abq88810 Human pro
c	28	16	55.2	863	4	AAH07593	Aah07593 Human cDN
c	29	16	55.2	991	4	AAK84716	Aak84716 Human imm
c	30	16	55.2	1166	6	ABX15078	Abx15078 cDNA enco
c	31	16	55.2	1201	10	ADC87218	Adc87218 Human GPC
c	32	16	55.2	1386	13	ADT48620	Adt48620 Bacterial
	33	16	55.2	1610	4	AAH15274	Aah15274 Human cDN
c	34	16	55.2	1638	3	AAC69417	Aac69417 Human sec
	35	16	55.2	1911	4	AAH14794	Aah14794 Human cDN
c	36	16	55.2	1989	10	ADC87088	Adc87088 Human GPC
	37	16	55.2	2103	10	ACC57312	Acc57312 Zinc fing
c	38	16	55.2	2159	12	ADQ64710	Adq64710 Novel hum
c	39	16	55.2	2756	4	AAH14861	Aah14861 Human cDN
c	40	16	55.2	3015	13	ADR07944	Adr07944 Full leng
	41	16	55.2	3187	10	ADB62850	Adb62850 Human cDN
	42	16	55.2	3608	4	AAK83192	Aak83192 Human imm
	43	16	55.2	3608	4	AAK74891	Aak74891 Human imm
c	44	16	55.2	3608	4	AAK67271	Aak67271 Human imm
	45	16	55.2	3791	5	AAF93891	Aaf93891 Human cDN
	46	16	55.2	3791	14	ADY63290	Ady63290 Human clo
	47	16	55.2	3810	11	ABD07975	Abd07975 Pseudomon
c	48	16	55.2	4428	11	ADI31783	Adi31783 Human cDN
c	49	16	55.2	4428	13	ADS83850	Ads83850 Human lym
c	50	16	55.2	4891	10	ADE71240	Ade71240 Novel hum
c	51	16	55.2	4891	10	ADF74248	Adf74248 Human nov
	52	16	55.2	5889	2	AAV31499	Aav31499 Murine mu
	53	16	55.2	5889	2	AAX21978	Aax21978 Mouse mul
	54	16	55.2	5889	2	AAX19819	Aax19819 Mouse mul
	55	16	55.2	5889	3	AAZ39557	Aaz39557 Murine mu
	56	16	55.2	5889	3	AAZ90194	Aaz90194 Murine mu
	57	16	55.2	7299	4	AAK77941	Aak77941 Human imm
	58	16	55.2	7299	5	ABA16794	Aba16794 Human ner
c	59	16	55.2	11234	5	ABA20857	Aba20857 Human ner
c	60	16	55.2	12123	4	AAK90226	Aak90226 Human dig
c	61	16	55.2	12123	5	AAS39855	Aas39855 Genomic s
c	62	16	55.2	12123	9	ADB32815	Adb32815 Human nov
	63	16	55.2	12394	4	AAD14749	Aad14749 Human gly
c	64	16	55.2	12542	4	AAS26800	Aas26800 Human gen
c	65	16	55.2	12542	8	ABX74149	Abx74149 Human nov
c	66	16	55.2	13255	4	AAK76842	Aak76842 Human imm
	67	16	55.2	13255	8	ADA41522	Ada41522 Human sec
	68	16	55.2	13255	10	ADA57647	Ada57647 BAC fragm
c	69	16	55.2	20001	14	AEB96548	Aeb96548 Human CD8
	70	16	55.2	23241	4	AAK87225	Aak87225 Human imm
	71	16	55.2	23241	4	AAK84291	Aak84291 Human imm
	72	16	55.2	23580	4	AAK83578	Aak83578 Human imm
c	73	16	55.2	23580	4	AAK66230	Aak66230 Human imm
	74	16	55.2	24934	4	AAK78433	Aak78433 Human imm
	75	16	55.2	24934	4	AAK71542	Aak71542 Human imm
	76	16	55.2	29829	11	ACN44310	Acn44310 Human gen
	77	16	55.2	30967	2	AAT32454	Aat32454 Calpain l
	78	16	55.2	32184	4	AAL04612	Aal04612 Human rep
	79	16	55.2	32184	4	ABL97535	Abl97535 Human tes
	80	16	55.2	32184	5	ABA18001	Aba18001 Human ner
	81	16	55.2	32194	4	AAL04340	Aal04340 Human rep
c	82	16	55.2	32367	2	AAV35620	Aav35620 Human SHO
c	83	16	55.2	33147	4	AAK67282	Aak67282 Human imm
c	84	16	55.2	37476	14	ADW38519	Adw38519 Immunomod
	85	16	55.2	39522	11	ACN44206	Acn44206 Human gen
c	86	16	55.2	41787	13	ABD33599	Abd33599 Human can
	87	16	55.2	42571	10	ABX14765	Abx14765 DNA encod
	88	16	55.2	42571	10	ACA61151	Aca61151 DNA encod
c	89	16	55.2	45698	12	ADJ12490	Adj12490 DNA fragm
c	90	16	55.2	46625	13	ABD33042	Abd33042 Mouse can
	91	16	55.2	53242	11	ACN43926	Acn43926 Human gen
	92	16	55.2	62124	13	ABD32754	Abd32754 Human can

	93	16	55.2	66973	11	ACN44230 Acn44230 Human gen
	94	16	55.2	76670	11	ACN45214 Acn45214 Human gen
c	95	16	55.2	93390	10	ADD71350 Add71350 Glutamine
c	96	16	55.2	97835	6	ABK84796 Abk84796 Human cDN
	97	16	55.2	110000	14	AEA61169_2 Continuation (3 of
c	98	16	55.2	110000	14	AEA61095_0 Aea61095 Human LOC
c	99	16	55.2	111084	12	ADQ18808 Adq18808 Human sof
	100	16	55.2	114771	12	ADQ17641 Adq17641 Human sof
c	101	16	55.2	125515	10	ADL13941 Adl13941 Osteoarth
c	102	16	55.2	127508	13	ABD33171 Abd33171 Human can
	103	16	55.2	128034	10	ADE43582 Ade43582 Polymorph
	104	16	55.2	128034	10	ADE43581 Ade43581 Human IDE
	105	16	55.2	128034	12	ADH54059 Adh54059 Human IDE
	106	16	55.2	128034	12	ADH54060 Adh54060 Human IDE
	107	16	55.2	130207	11	ACN44762 Acn44762 Human gen
c	108	16	55.2	131673	12	ADQ21602 Adq21602 Human sof
	109	16	55.2	142318	11	ACN44850 Acn44850 Human gen
c	110	16	55.2	144723	11	ACN44898 Acn44898 Human gen
	111	16	55.2	152330	11	ACN45070 Acn45070 Human gen
c	112	16	55.2	153752	12	ADQ97531 Adq97531 Human can
	113	16	55.2	160274	14	AEB32377 Aeb32377 Human gen
	114	16	55.2	160300	14	AEB32388 Aeb32388 Human gen
	115	16	55.2	162450	3	AAZ86967 Aaz86967 Retinobla
	116	16	55.2	165199	6	ABK83460 Abk83460 Human cDN
c	117	16	55.2	166181	12	ADQ20461 Adq20461 Human sof
c	118	16	55.2	166181	12	ADQ18633 Adq18633 Human sof
	119	16	55.2	168174	6	ABT11173 Abt11173 Human 5-l
	120	16	55.2	168273	6	ABT11114 Abt11114 Human 5-l
	121	16	55.2	177866	10	ADL13935 Adl13935 Osteoarth
c	122	16	55.2	198522	11	ACN44010 Acn44010 Human gen
c	123	16	55.2	202100	10	ADE43315 Ade43315 Human IDE
c	124	16	55.2	202100	12	ADH54357 Adh54357 Human IDE
	125	16	55.2	203654	10	ABX16034 Abx16034 Human gen
c	126	16	55.2	215126	12	ADQ97362 Adq97362 Mouse can
c	127	16	55.2	256190	13	ABD33276 Abd33276 Human can
	128	16	55.2	260209	6	ABS56564 Abs56564 Human SUL
	129	16	55.2	260209	12	ADN16204 Adn16204 Human sul
c	130	16	55.2	310268	13	ABD32548 Abd32548 Human can
	131	16	55.2	337344	13	ABD32715 Abd32715 Human can
	132	16	55.2	344548	11	ACN44070 Acn44070 Human gen
c	133	15	51.7	265	6	ABN23484 Abn23484 Human ORF
c	134	15	51.7	320	8	AAD49104 Aad49104 Mouse DST
c	135	15	51.7	327	4	AAK75983 Aak75983 Human imm
c	136	15	51.7	327	4	AAK75984 Aak75984 Human imm
c	137	15	51.7	328	4	AAK75985 Aak75985 Human imm
	138	15	51.7	333	10	ADG91042 Adg91042 Hepatic s
c	139	15	51.7	336	6	ABL84872 Abl84872 Human ova
	140	15	51.7	345	13	ADS14740 Ads14740 Pseudomon
c	141	15	51.7	370	5	AAF65791 Aaf65791 Novel hum
c	142	15	51.7	389	3	AAH30877 Aah30877 Human col
	143	15	51.7	392	3	AAA57072 Aaa57072 Human col
	144	15	51.7	392	6	ABT12494 Abt12494 Orestes s
	145	15	51.7	392	10	ACD91788 Acd91788 Human col
	146	15	51.7	414	11	ABD14593 Abd14593 Pseudomon
	147	15	51.7	435	10	ADG91043 Adg91043 Hepatic s
c	148	15	51.7	481	10	ADK53699 Adk53699 Plant DNA
c	149	15	51.7	492	5	ABV18599 Abv18599 Human pro
	150	15	51.7	493	3	AAA57090 Aaa57090 Human col
	151	15	51.7	493	6	ABT12512 Abt12512 Orestes s
	152	15	51.7	493	10	ACD91806 Acd91806 Human col
c	153	15	51.7	500	3	AAC52805 Aac52805 Arabidops
	154	15	51.7	512	4	AAI83598 Aai83598 Human pol
	155	15	51.7	518	4	AAL34934 Aal34934 Human mus
	156	15	51.7	518	8	ABX57922 Abx57922 cDNA enco
	157	15	51.7	518	12	ADJ27649 Adj27649 Human mus
	158	15	51.7	519	14	ACL57426 Acl57426 Human col
c	159	15	51.7	521	5	ADI68350 Adi68350 Human ova
c	160	15	51.7	521	5	ADI74716 Adi74716 Human ova
	161	15	51.7	528	11	ABD11716 Abd11716 Pseudomon
c	162	15	51.7	531	13	ACF85251 Acf85251 Human SIR
c	163	15	51.7	535	5	ADL39953 Adl39953 Human ova
	164	15	51.7	549	4	AAH11651 Aah11651 Human cDN
c	165	15	51.7	563	12	ADN12912 Adn12912 Human pro
c	166	15	51.7	572	5	ABV48382 Abv48382 Human pro
	167	15	51.7	601	5	ABV54558 Abv54558 Human pro
	168	15	51.7	622	3	AAA79480 Aaa79480 Eucalyptu
c	169	15	51.7	632	10	ADK55993 Adk55993 Plant DNA
	170	15	51.7	652	10	ADD34348 Add34348 Mouse mit
c	171	15	51.7	730	4	AAI96902 Aai96902 Human neu
c	172	15	51.7	801	5	ADL45827 Adl45827 Human ova
c	173	15	51.7	865	10	ADI62659 Adi62659 Human apo
	174	15	51.7	896	10	ADD55880 Add55880 Thalecres
	175	15	51.7	909	3	AAC35208 Aac35208 Arabidops
	176	15	51.7	914	12	ADO01642 Ado01642 Thalecres
	177	15	51.7	914	12	ADO61552 Ado61552 Transcrip
	178	15	51.7	947	3	AAC39436 Aac39436 Arabidops
c	179	15	51.7	1012	6	ABK92489 Abk92489 Human pro
c	180	15	51.7	1372	10	ADB99859 Adb99859 Mouse alp
	181	15	51.7	1389	3	AAC79822 Aac79822 Human sec
c	182	15	51.7	1472	6	ABK86368 Abk86368 Secretory
c	183	15	51.7	1472	11	ADP65639 Adp65639 Human sec
c	184	15	51.7	1563	9	ADA31544 Ada31544 DNA encod
	185	15	51.7	1600	2	AAV69205 Aav69205 Sequence
	186	15	51.7	1600	3	AAZ24361 Aaz24361 Human ICA
	187	15	51.7	1600	6	ABK09378 Abk09378 Intercell
	188	15	51.7	1600	10	ADG25767 Adg25767 Human ICA
c	189	15	51.7	1613	10	ADC64873 Adc64873 Human bet
c	190	15	51.7	1613	10	ADG25197 Adg25197 Human GP3
c	191	15	51.7	1735	6	ABK73679 Abk73679 Bacillus
c	192	15	51.7	1892	6	ABK97371 Abk97371 Secretory
c	193	15	51.7	2053	5	AAS89320 Aas89320 DNA encod
c	194	15	51.7	2121	4	AAH16285 Aah16285 Human cDN
	195	15	51.7	2234	5	AAS69659 Aas69659 DNA encod
	196	15	51.7	2402	5	AAS76449 Aas76449 DNA encod
c	197	15	51.7	3093	13	ADU50860 Adu50860 Human thi
c	198	15	51.7	3131	2	AAV40561 Aav40561 Human thi
	199	15	51.7	3131	5	AAS77500 Aas77500 DNA encod
c	200	15	51.7	3131	14	ADZ49180 Adz49180 Insulin s
	201	15	51.7	3218	12	ADQ63784 Adq63784 Novel hum
c	202	15	51.7	3492	11	ADM03005 Adm03005 Human cDN
c	203	15	51.7	3695	6	ABS76539 Abs76539 cDNA enco
	204	15	51.7	3715	3	AAA79735 Aaa79735 Eucalyptu
	205	15	51.7	3728	8	ABZ36216 Abz36216 Human sec
	206	15	51.7	3747	12	ADQ65013 Adq65013 Novel hum
c	207	15	51.7	4211	6	ABL68268 Abl68268 Kidney ca
c	208	15	51.7	4211	6	ABL68580 Abl68580 Kidney ca
c	209	15	51.7	4211	6	ABL68879 Abl68879 Kidney ca
c	210	15	51.7	4280	4	AAL06145 Aal06145 Human rep
c	211	15	51.7	4280	4	ABL98710 Abl98710 Human tes
c	212	15	51.7	4283	4	AAL06146 Aal06146 Human rep
c	213	15	51.7	4283	4	ABL98711 Abl98711 Human tes
	214	15	51.7	4621	4	ABA05827 Aba05827 Human Lys
	215	15	51.7	4978	4	AAK89129 Aak89129 Human dig
	216	15	51.7	4978	5	AAS31879 Aas31879 Human liv
	217	15	51.7	4978	6	ABN90234 Abn90234 Human liv
	218	15	51.7	4978	11	ADJ15147 Adj15147 Human liv
	219	15	51.7	5757	5	AAS74316 Aas74316 DNA encod
	220	15	51.7	6430	2	AAX02992 Aax02992 Human IL-
c	221	15	51.7	7943	4	AAK83418 Aak83418 Human imm
	222	15	51.7	11101	6	ABN83947 Abn83947 Human tra
c	223	15	51.7	12292	4	AAS59537 Aas59537 Propionib
c	224	15	51.7	12292	8	ACF64466 Acf64466 Propionib
	225	15	51.7	13630	4	AAS27709 Aas27709 DNA encod
	226	15	51.7	13630	4	ABK44032 Abk44032 Genomic D
	227	15	51.7	13630	10	ADB94512 Adb94512 Novel hum
	228	15	51.7	13630	13	ADI55017 Adi55017 Novel hum
	229	15	51.7	14769	4	AAL04404 Aal04404 Human rep
	230	15	51.7	14781	4	AAK66710 Aak66710 Human imm
c	231	15	51.7	15914	4	AAK84889 Aak84889 Human imm
	232	15	51.7	15914	8	ABZ73768 Abz73768 Secreted
	233	15	51.7	15914	8	ADA98447 Ada98447 Human sec
	234	15	51.7	15914	8	ADA44234 Ada44234 Human sec
	235	15	51.7	15914	10	ADC20616 Adc20616 Human sec
	236	15	51.7	15914	10	ADF10823 Adf10823 Human sec
	237	15	51.7	15914	10	ABT16886 Abt16886 Human sec
	238	15	51.7	15914	10	ABZ67365 Abz67365 Human sec

239	15	51.7	17483	13	ADT89185	Adt89185 Human uri
c 240	15	51.7	17752	6	AAD45439	Aad45439 Human dru
c 241	15	51.7	17752	12	ADL67252	Adl67252 Human dru
242	15	51.7	18887	6	AAS18543	Aas18543 DNA encod
c 243	15	51.7	19371	4	AAL05706	Aal05706 Human rep
244	15	51.7	19533	10	ADC87260	Adc87260 Human GPC
245	15	51.7	19851	4	AAK85642	Aak85642 Human imm
246	15	51.7	19929	4	AAL07279	Aal07279 Human rep
247	15	51.7	19929	4	ABL98825	Abl98825 Human tes
248	15	51.7	20907	4	AAL07278	Aal07278 Human rep
249	15	51.7	20907	4	ABL98824	Abl98824 Human tes
c 250	15	51.7	22324	12	ADQ59488	Adq59488 Human can
c 251	15	51.7	22328	14	ADZ13857	Adz13857 Human can
c 252	15	51.7	27154	4	AAL05708	Aal05708 Human rep
253	15	51.7	32195	4	ABK42625	Abk42625 Genomic s
254	15	51.7	32195	9	ADB60781	Adb60781 Connectiv
255	15	51.7	34261	13	ABD33518	Abd33518 Murine ca
256	15	51.7	34745	14	ADZ13487	Adz13487 Murine ca
c 257	15	51.7	43058	6	ABL64982	Abl64982 Lung canc
c 258	15	51.7	43058	6	ABL65219	Abl65219 Lung canc
c 259	15	51.7	43058	6	ABN97455	Abn97455 Gene #395
260	15	51.7	43795	3	AAZ92583	Aaz92583 Human DAZ
261	15	51.7	46050	10	ABX13974	Abx13974 Human Ras
262	15	51.7	46718	8	ABS57422	Abs57422 Human pro
263	15	51.7	46718	10	ADG62974	Adg62974 Human pro
c 264	15	51.7	48208	13	ADU50859	Adu50859 Human thi
265	15	51.7	49243	4	ABL03188	Abl03188 Drosophil
266	15	51.7	51558	13	ACN37207	Acn37207 Human per
267	15	51.7	51935	4	AAK75883	Aak75883 Human imm
c 268	15	51.7	52287	14	ADZ42275	Adz42275 Human thi
269	15	51.7	53332	6	AAL48890	Aal48890 Human Pft
c 270	15	51.7	53795	9	ADA02858	Ada02858 Human LMO
c 271	15	51.7	53795	10	ADB72596	Adb72596 Human LMO
c 272	15	51.7	53795	10	ADC85337	Adc85337 Mouse Lmo
c 273	15	51.7	53795	12	ADM74453	Adm74453 Human car
c 274	15	51.7	54000	8	AAD52261	Aad52261 Human int
c 275	15	51.7	54000	14	AEB44785	Aeb44785 Human int
276	15	51.7	54592	14	AEA61160_3	Continuation (4 of
277	15	51.7	57652	12	ADQ59398_3	Continuation (4 of
278	15	51.7	69652	13	ABD33115	Abd33115 Human can
279	15	51.7	76138	12	ADQ97334	Adq97334 Mouse can
280	15	51.7	78361	10	AAL52246_3	Continuation (4 of
281	15	51.7	83120	9	AAL57571	Aal57571 Human CGI
c 282	15	51.7	83400	12	ADP07906	Adp07906 Human RAD
283	15	51.7	84539	6	ABL64158	Abl64158 Stomach c
284	15	51.7	84539	10	ADL13479	Adl13479 Osteoarth
285	15	51.7	88939	13	ABD32535_3	Continuation (4 of
286	15	51.7	88939	14	ADZ13035_3	Continuation (4 of
287	15	51.7	89182	11	ACN44582_3	Continuation (4 of
c 288	15	51.7	90442	9	ADA03077	Ada03077 Mouse mCG
c 289	15	51.7	90442	9	ADA66361	Ada66361 Mouse mCG
c 290	15	51.7	90442	10	ADB72815	Adb72815 Mouse mCG
c 291	15	51.7	90442	10	ADC26997	Adc26997 Mouse car
c 292	15	51.7	90442	11	ADL27155	Adl27155 Mouse gen
c 293	15	51.7	95484	12	ADQ97298	Adq97298 Mouse can
294	15	51.7	110000	4	AAK95240_06	Continuation (7 of
295	15	51.7	110000	4	AAK96733_06	Continuation (7 of
296	15	51.7	110000	6	ABS55320_2	Continuation (3 of
297	15	51.7	110000	6	ABT00010_06	Continuation (7 of
298	15	51.7	110000	6	ABT01503_06	Continuation (7 of
299	15	51.7	110000	11	ADW70291_06	Continuation (7 of
300	15	51.7	110000	12	ADH77486_06	Continuation (7 of
301	15	51.7	110000	12	ADQ59398_2	Continuation (3 of
302	15	51.7	110000	14	ADZ13665_2	Continuation (3 of
303	15	51.7	110000	14	ADZ13665_3	Continuation (4 of
304	15	51.7	110000	14	AEA61163_2	Continuation (3 of
305	15	51.7	113000	9	ABT44365	Abt44365 Partial g
c 306	15	51.7	116592	8	ABX15519	Abx15519 Human tyr
c 307	15	51.7	116592	10	AAD47900	Aad47900 Human tra
c 308	15	51.7	116592	14	AEB47448	Aeb47448 Human sul
309	15	51.7	121160	12	ADQ97870	Adq97870 Human can
310	15	51.7	134738	11	ACN44182	Acn44182 Human gen
311	15	51.7	139257	10	ADC89520	Adc89520 Human COR
c 312	15	51.7	147300	12	ADP45593	Adp45593 Human Rho
c 313	15	51.7	147700	14	ADX98570	Adx98570 Human gua
314	15	51.7	156652	14	AEA61138	Aea61138 Human UGT
315	15	51.7	168828	14	ADZ13592	Adz13592 Human can
316	15	51.7	175338	11	ACN45088	Acn45088 Mouse gen
c 317	15	51.7	178024	12	ADQ97721	Adq97721 Human can
c 318	15	51.7	178896	6	ABQ88146	Abq88146 Human ost
c 319	15	51.7	215248	12	ADQ97284	Adq97284 Mouse can
320	15	51.7	224112	13	ABD32600	Abd32600 Mouse can
321	15	51.7	240000	8	ACD13446	Acd13446 Human DNA
322	15	51.7	243390	13	ABD33366	Abd33366 Human can
323	15	51.7	243934	14	ADZ13446	Adz13446 Human can
c 324	15	51.7	264965	12	ADN16203	Adn16203 Human sul
c 325	15	51.7	268685	6	ABS56563	Abs56563 Human SUL
326	15	51.7	272022	12	ADQ97126	Adq97126 Human can
327	15	51.7	325791	4	AAS43104	Aas43104 Human Oes
c 328	15	51.7	325791	4	AAS43104	Aas43104 Human Oes
329	15	51.7	341511	6	ABS55200	Abs55200 Genomic D

ALIGNMENTS

```
RESULT 1
AAS15366
ID   AAS15366 standard; cDNA; 2145 BP.
XX
AC   AAS15366;
XX
DT   16-JAN-2002  (first entry)
XX
DE   cDNA encoding human PRO9964 polypeptide.
XX
KW   Human; PRO9964; clone DNA96973; immune-related disorder;
KW   inflammatory disorder; infectious disorder; immunodeficiency disorder;
KW   autoimmune disorder; renal disease; demyelinating disease; skin disease;
KW   neoplasia; transplantation associated disease; gene therapy;
KW   immunosuppressive; anti-inflammatory; antiasthmatic; antidiabetic; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             87. .1601
FT                   /*tag= a
FT                   /product= "PRO9964 polypeptide"
XX
PN   WO200166740-A2.
XX
PD   13-SEP-2001.
XX
PF   01-MAR-2001; 2001WO-US006666.
XX
PR   03-MAR-2000; 2000US-0187202P.
PR   21-MAR-2000; 2000US-0191015P.
PR   30-MAY-2000; 2000WO-US014941.
PR   05-JUN-2000; 2000US-0209832P.
PR   24-AUG-2000; 2000WO-US023328.
PR   01-DEC-2000; 2000WO-US032678.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Eaton DL,  Fong S,  Goddard A,  Godowski PJ,  Grimaldi CJ,  Gurney AL;
PI   Tumas D,  Watanabe CK,  Wood WI,  Zhang Z;
XX
DR   WPI; 2001-625876/72.
DR   P-PSDB; AAU09184.
XX
PT   Nucleic acids encoding PRO polypeptides, useful for detecting and
PT   treating immune related diseases and disorders in mammals including
PT   autoimmune diseases, inflammatory diseases and asthma.
XX
PS   Claim 2; Fig 13; 122pp; English.
XX
```

```
CC   The present invention relates to the isolation of 9 novel human PRO
CC   polypeptides (AAU09178-AAU09186) and the cDNA sequences encoding them.
CC   The novel PRO polypeptides include PRO1356, PRO1268, PRO1884, PRO3444,
CC   PRO3151, PRO4322, PRO9964, PRO10008 and PRO19598. The cDNA sequences
CC   encoding these PRO polypeptides have been designated as clones DNA64886-
CC   1601, DNA64903-1553, DNA84318-2520, DNA87997, DNA89273, DNA92223-2567,
CC   DNA96973, DNA101921 and DNA145887 respectively. Compositions (e.g.
CC   vaccines) containing PRO polypeptides and methods of using these
CC   compositions are useful in the treatment and diagnosis of immune-related
CC   disorders. Such disorders include immune-mediated inflammatory disorders
CC   (e.g. osteoarthritis), non-immune-mediated inflammatory disorders (e.g.
CC   diabetes mellitus), infectious disorders (e.g. granulomatous hepatitis),
CC   immunodeficiency disorders (e.g. AIDS), autoimmune disorders (e.g.
CC   rheumatoid arthritis), immune-related renal diseases (e.g. cirrhosis),
CC   demyelinating diseases of the peripheral or central nervous system (e.g.
CC   Guillain-Barre syndrome), immune-mediated skin diseases (e.g. contact
CC   dermatitis), neoplasias and transplantation associated diseases. The
CC   polynucleotide sequences of the invention may be used in gene therapy.
CC   AAS15360-AAS15368 represent cDNA sequences encoding for the novel human
CC   PRO polypeptides of the invention
XX
SQ   Sequence 2145 BP; 608 A; 541 C; 551 G; 445 T; 0 U; 0 Other;

  Query Match             65.5%;  Score 19;  DB 4;  Length 2145;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 GCTCAGGCATGAGCCAGCA 29
              |||||||||||||||||||
Db          1 GCTCAGGCATGAGCCAGCA 19


RESULT 2
ADQ97385/c
ID   ADQ97385 standard; DNA; 21728 BP.
XX
AC   ADQ97385;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Human cancer associated sequence HD08-041, SEQ ID 362.
XX
KW   Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.
XX
OS   Homo sapiens.
XX
PN   WO2004060304-A2.
XX
PD   22-JUL-2004.
XX
PF   22-DEC-2003; 2003WO-US041389.
XX
PR   27-DEC-2002; 2002US-00330773.
XX
PA   (SAGR-) SAGRES DISCOVERY INC.
XX
PI   Morris DW,  Malandro MS;
XX
DR   WPI; 2004-543781/52.
XX
PT   New isolated cancer associated nucleic acids comprising at least 10
PT   contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT   cancers such as leukemia and lymphoma.
XX
PS   Claim 1; SEQ ID NO 362; 199pp; English.
XX
CC   The present invention relates to cancer associated sequences (ADQ97025-
CC   ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC   treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic formate directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 21728 BP; 3820 A; 6429 C; 6469 G; 4239 T; 0 U; 771 Other;

  Query Match             58.6%;  Score 17;  DB 12;  Length 21728;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAGG 17
              |||||||||||||||||
Db       5609 CCTCTCTGGAGCTCAGG 5593


RESULT 3
ABV56351/c
ID   ABV56351 standard; cDNA; 391 BP.
XX
AC   ABV56351;
XX
DT   17-SEP-2002  (first entry)
XX
DE   Human prostate expression marker cDNA 56342.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US005171.
XX
PR   17-FEB-2000; 2000US-0183319P.
PR   16-MAR-2000; 2000US-0189862P.
PR   25-MAY-2000; 2000US-0207454P.
PR   09-JUN-2000; 2000US-0211314P.
PR   18-JUL-2000; 2000US-0219007P.
PR   13-DEC-2000; 2000US-0255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Endege WO,  Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS   Claim 1; Page 10875-10876; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement. (I) is useful for: (a) assessing whether
CC   a patient is afflicted with prostate cancer; (b) monitoring the
CC   progression of prostate cancer in a patient; (c) assessing the efficacy
CC   of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC   the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC   determining whether prostate cancer has metastasized in a patient; (h)
CC   assessing the aggressiveness or indolence of prostate cancer in a patient
CC   ; (I) is also useful as a pharmacodyanamic or pharmacogenomic marker
XX
SQ   Sequence 391 BP; 117 A; 89 C; 86 G; 99 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 5;  Length 391;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db         285 CAGGCATGAGCCAGCA 270
```

```
RESULT 4
ACH17384
ID   ACH17384 standard; cDNA; 394 BP.
XX
AC   ACH17384;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human adult heart cDNA #1698.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 4596; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 394 BP; 84 A; 102 C; 90 G; 118 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 9;  Length 394;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        175 CAGGCATGAGCCAGCA 190


RESULT 5
ABV58668
ID   ABV58668 standard; cDNA; 410 BP.
XX
AC   ABV58668;
XX
DT   13-SEP-2002  (first entry)
XX
DE   Human prostate expression marker cDNA 58659.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US005171.
XX
PR   17-FEB-2000; 2000US-0183319P.
PR   16-MAR-2000; 2000US-0189862P.
PR   25-MAY-2000; 2000US-0207454P.
PR   09-JUN-2000; 2000US-0211314P.
PR   18-JUL-2000; 2000US-0219007P.
PR   13-DEC-2000; 2000US-0255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Endege WO,  Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS   Claim 1; Page 11256; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement. (I) is useful for: (a) assessing whether
CC   a patient is afflicted with prostate cancer; (b) monitoring the
CC   progression of prostate cancer in a patient; (c) assessing the efficacy
CC   of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC   the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC   determining whether prostate cancer has metastasized in a patient; (h)
CC   assessing the aggressiveness or indolence of prostate cancer in a patient
CC   ; (I) is also useful as a pharmacodyanamic or pharmacogenomic marker
XX
SQ   Sequence 410 BP; 136 A; 86 C; 74 G; 114 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 5;  Length 410;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        258 CAGGCATGAGCCAGCA 273


RESULT 6
ABV18799
ID   ABV18799 standard; cDNA; 426 BP.
XX
AC   ABV18799;
XX
DT   13-SEP-2002  (first entry)
XX
DE   Human prostate expression marker cDNA 18790.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
```

```
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US005171.
XX
PR   17-FEB-2000; 2000US-0183319P.
PR   16-MAR-2000; 2000US-0189862P.
PR   25-MAY-2000; 2000US-0207454P.
PR   09-JUN-2000; 2000US-0211314P.
PR   18-JUL-2000; 2000US-0219007P.
PR   13-DEC-2000; 2000US-0255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Endege WO,  Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS   Claim 1; Page 3091; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement. (I) is useful for: (a) assessing whether
CC   a patient is afflicted with prostate cancer; (b) monitoring the
CC   progression of prostate cancer in a patient; (c) assessing the efficacy
CC   of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC   the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC   determining whether prostate cancer has metastasized in a patient; (h)
CC   assessing the aggressiveness or indolence of prostate cancer in a patient
CC   ; (I) is also useful as a pharmacodyanamic or pharmacogenomic marker
XX
SQ   Sequence 426 BP; 143 A; 83 C; 78 G; 116 T; 0 U; 6 Other;

  Query Match             55.2%;  Score 16;  DB 5;  Length 426;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        248 CAGGCATGAGCCAGCA 263


RESULT 7
ACH14691/c
ID   ACH14691 standard; cDNA; 429 BP.
XX
AC   ACH14691;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human adult brain cDNA #1903.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 1903; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 429 BP; 84 A; 124 C; 132 G; 85 T; 0 U; 4 Other;

  Query Match             55.2%;  Score 16;  DB 9;  Length 429;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        420 CAGGCATGAGCCAGCA 405


RESULT 8
ACH25630/c
ID   ACH25630 standard; cDNA; 443 BP.
XX
AC   ACH25630;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human adult ovary cDNA #4010.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
```

PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
DR WPI; 2003-615964/58.
XX
PT New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
PS Claim 1; SEQ ID NO 12842; 44pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversities, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ Sequence 443 BP; 81 A; 138 C; 131 G; 92 T; 0 U; 1 Other;

Query Match 55.2%; Score 16; DB 9; Length 443;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 CAGGCATGAGCCAGCA 29
||||||||||||||||
Db 413 CAGGCATGAGCCAGCA 398

RESULT 9
ACH24672/c
ID ACH24672 standard; cDNA; 450 BP.
XX
AC ACH24672;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human adult ovary cDNA #3052.
XX
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
DR WPI; 2003-615964/58.
XX
PT New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
PS Claim 1; SEQ ID NO 11884; 44pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversities, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ Sequence 450 BP; 151 A; 96 C; 91 G; 103 T; 0 U; 9 Other;

Query Match 55.2%; Score 16; DB 9; Length 450;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 CAGGCATGAGCCAGCA 29
||||||||||||||||
Db 231 CAGGCATGAGCCAGCA 216

RESULT 10
ABN18360/c
ID ABN18360 standard; cDNA; 452 BP.
XX
AC ABN18360;
XX
DT 24-JUN-2002 (first entry)
XX
DE Human ORFX polynucleotide sequence SEQ ID NO:5197.
XX
KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192523-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US010836.
XX
PR 30-MAY-2000; 2000US-0206132P.
PR 29-AUG-2000; 2000US-0228716P.
XX
PA (CURA-) CURAGEN CORP.
XX

```
PI   Shimkets RA,  Leach MD;
XX
DR   WPI; 2002-106308/14.
DR   P-PSDB; ABP02608.
XX
PT   Novel human polypeptides and polynucleotides useful for diagnosing,
PT   preventing and treating cardiovascular disease, neurodegenerative,
PT   hyperproliferative disorders and autoimmune disorders.
XX
PS   Disclosure; SEQ ID NO 5197; 1037pp; English.
XX
CC   The present invention describes substantially purified human proteins
CC   (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC   in the specification). ABN15762 to ABN27252 encode the human ORFX
CC   proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC   treating or preventing a pathology associated with an ORFX-associated
CC   disorder in humans, and in the manufacture of a medicament for treating a
CC   syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC   sequences can be used in gene therapy. ORFX sequences can be used in the
CC   treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC   psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC   osteoarthritis, neurodegenerative disorders, disorders related to organ
CC   transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC   lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC   storage disease, various immune deficiencies and disorders, infectious
CC   diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC   arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC   disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC   useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC   bone degenerative disorders, or periodontal disease, and for gut
CC   protection or regeneration and treatment of lung or liver fibrosis,
CC   reperfusion injury in various tissues and conditions resulting from
CC   systemic cytokine damage. N.B. The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 452 BP; 120 A; 115 C; 100 G; 116 T; 0 U; 1 Other;

  Query Match             55.2%;  Score 16;  DB 6;  Length 452;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAG 16
              ||||||||||||||||
Db        341 CCTCTCTGGAGCTCAG 326


RESULT 11
AAI12911
ID   AAI12911 standard; DNA; 463 BP.
XX
AC   AAI12911;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #2844 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 2844; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 463 BP; 126 A; 95 C; 160 G; 82 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 4;  Length 463;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAG 23
              ||||||||||||||||
Db         36 GGAGCTCAGGCATGAG 51


RESULT 12
ABA54612
ID   ABA54612 standard; DNA; 463 BP.
XX
AC   ABA54612;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #2917.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
```

```
PS   Claim 1; SEQ ID NO 2917; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 463 BP; 126 A; 95 C; 160 G; 82 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 4;  Length 463;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAG 23
              ||||||||||||||||
Db         36 GGAGCTCAGGCATGAG 51


RESULT 13
AAI34271
ID   AAI34271 standard; DNA; 463 BP.
XX
AC   AAI34271;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #2957 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 2957; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 463 BP; 126 A; 95 C; 160 G; 82 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 4;  Length 463;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAG 23
              ||||||||||||||||
Db         36 GGAGCTCAGGCATGAG 51


RESULT 14
ABA44163
ID   ABA44163 standard; DNA; 463 BP.
XX
AC   ABA44163;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #2858.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 1; SEQ ID NO 2858; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 463 BP; 126 A; 95 C; 160 G; 82 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 4;  Length 463;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAG 23
```

```
                  ||||||||||||||||
Db          36 GGAGCTCAGGCATGAG 51


RESULT 15
ABA24397
ID   ABA24397 standard; DNA; 463 BP.
XX
AC   ABA24397;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #2863 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000666.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts.
XX
PS   Claim 1; SEQ ID NO 2863; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 463 BP; 126 A; 95 C; 160 G; 82 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 4;  Length 463;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAG 23
              ||||||||||||||||
Db         36 GGAGCTCAGGCATGAG 51


RESULT 16
AAK28345
ID   AAK28345 standard; DNA; 463 BP.
XX
AC   AAK28345;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 2902.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 2902; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 463 BP; 126 A; 95 C; 160 G; 82 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 4;  Length 463;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAG 23
              ||||||||||||||||
Db         36 GGAGCTCAGGCATGAG 51


RESULT 17
AAK02903
ID   AAK02903 standard; DNA; 463 BP.
XX
AC   AAK02903;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 2894.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
```

PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 2894; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention
XX
SQ Sequence 463 BP; 126 A; 95 C; 160 G; 82 T; 0 U; 0 Other;

Query Match 55.2%; Score 16; DB 4; Length 463;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GGAGCTCAGGCATGAG 23
||||||||||||||||
Db 36 GGAGCTCAGGCATGAG 51

RESULT 18
ABS27944
ID ABS27944 standard; DNA; 463 BP.
XX
AC ABS27944;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID No 2934.
XX
KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 1; SEQ ID NO 2934; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 463 BP; 126 A; 95 C; 160 G; 82 T; 0 U; 0 Other;

Query Match 55.2%; Score 16; DB 4; Length 463;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GGAGCTCAGGCATGAG 23
||||||||||||||||
Db 36 GGAGCTCAGGCATGAG 51

RESULT 19
AAI02830
ID AAI02830 standard; DNA; 463 BP.
XX
AC AAI02830;
XX
DT 09-OCT-2001 (first entry)
XX
DE Probe #2821 used to measure gene expression in human breast sample.
XX
KW Probe; human; breast disease; breast cancer; development disorder; ss;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US000661.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-476286/51.
XX
PT Novel single exon nucleic acid probe used to measuring gene expression in
PT a human breast.
XX
PS Claim 25; SEQ ID NO 2821; 322pp; English.

```
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and non-
CC   carcinoma tumours. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 463 BP; 126 A; 95 C; 160 G; 82 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 5;  Length 463;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAG 23
              ||||||||||||||||
Db         36 GGAGCTCAGGCATGAG 51


RESULT 20
ABS02854
ID   ABS02854 standard; DNA; 463 BP.
XX
AC   ABS02854;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe from lung SEQ ID No 2845.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 1; SEQ ID NO 2845; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 463 BP; 126 A; 95 C; 160 G; 82 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 6;  Length 463;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAG 23
              ||||||||||||||||
Db         36 GGAGCTCAGGCATGAG 51


RESULT 21
ABV50399/c
ID   ABV50399 standard; cDNA; 495 BP.
XX
AC   ABV50399;
XX
DT   17-SEP-2002  (first entry)
XX
DE   Human prostate expression marker cDNA 50390.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US005171.
XX
PR   17-FEB-2000; 2000US-0183319P.
PR   16-MAR-2000; 2000US-0189862P.
PR   25-MAY-2000; 2000US-0207454P.
```

PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 9809; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodyanamic or pharmacogenomic marker
XX
SQ Sequence 495 BP; 164 A; 105 C; 128 G; 98 T; 0 U; 0 Other;

Query Match 55.2%; Score 16; DB 5; Length 495;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 CAGGCATGAGCCAGCA 29
||||||||||||||||
Db 225 CAGGCATGAGCCAGCA 210

RESULT 22
ABV48581
ID ABV48581 standard; cDNA; 564 BP.
XX
AC ABV48581;
XX
DT 17-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 48572.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 9520; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodyanamic or pharmacogenomic marker
XX
SQ Sequence 564 BP; 176 A; 125 C; 113 G; 144 T; 0 U; 6 Other;

Query Match 55.2%; Score 16; DB 5; Length 564;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 CAGGCATGAGCCAGCA 29
||||||||||||||||
Db 307 CAGGCATGAGCCAGCA 322

RESULT 23
ABN63733/c
ID ABN63733 standard; cDNA; 573 BP.
XX
AC ABN63733;
XX
DT 28-JUN-2002 (first entry)
XX
DE Human cancer related polynucleotide SEQ ID NO 3700.
XX
KW Human; cytostatic; gene expression; gene mapping; tissue profiling;
KW gene therapy; cancer; tumour; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200214500-A2.
XX
PD 21-FEB-2002.
XX
PF 16-AUG-2001; 2001WO-US025840.
XX
PR 16-AUG-2000; 2000US-0226326P.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;
PI Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;
XX
DR WPI; 2002-241905/29.
XX
PT New nucleic acid for producing a polypeptide, detecting differentially
PT expressed genes correlated with a cancerous state of a mammalian cell,
PT and inhibiting tumor growth.
XX
PS Claim 1; SEQ ID NO 3700; 883pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
CC with cytostatic activity. The polynucleotide is used to produce a
CC polypeptide, to detect differentially expressed genes correlated with a
CC cancerous state of a mammalian cell and to inhibit tumour growth. The
CC polynucleotide is used as a probe in mapping and tissue profiling. The

CC encoded polypeptide and antibodies to the polypeptide can also be used
CC for therapeutic and diagnostic purposes. The polynucleotide is useful for
CC gene therapy. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 573 BP; 141 A; 170 C; 108 G; 154 T; 0 U; 0 Other;

Query Match 55.2%; Score 16; DB 6; Length 573;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 CAGGCATGAGCCAGCA 29
||||||||||||||||
Db 561 CAGGCATGAGCCAGCA 546

RESULT 24
AEB33143
ID AEB33143 standard; DNA; 601 BP.
XX
AC AEB33143;
XX
DT 08-SEP-2005 (first entry)
XX
DE Human DNA polymorphic region #723.
XX
KW SNP detection; diagnosis; non-insulin dependent diabetes; obesity;
KW antidiabetic; anorectic; endocrine disease; gastrointestinal disease;
KW metabolic disorder; nutritional disorder; single nucleotide polymorphism;
KW SNP; ds.
XX
OS Homo sapiens.
XX
PN US2005147987-A1.
XX
PD 07-JUL-2005.
XX
PF 19-JUL-2004; 2004US-00893315.
XX
PR 08-SEP-2000; 2000US-0231397P.
PR 10-SEP-2001; 2001US-00948947.
XX
PA (APPL-) APPLERA CORP NY.
XX
PI Venter JC, Zhang JN, Liu X, Rowe W, Cravchik A, Kalush F;
PI Naik A, Subramanian G, Woodage T;
XX
DR WPI; 2005-511776/52.
XX
PT New detection reagent capable of detecting 1, 100, 500, 1000 or 5000 or
PT more single nucleic acid polymorphisms, useful in identifying an
PT individual having or at risk of developing type II diabetes or obesity.
XX
PS Claim 13; SEQ ID NO 906; 31pp; English.
XX
CC The invention relates to a detection reagent capable of detecting one or
CC more single nucleic acid polymorphisms. The invention also relates to
CC determining whether a trait is linked to one of the human chromosomes or
CC its sub-region, a computer readable medium having stored in it the SNP
CC relational information given in the specification, an isolated nucleic
CC acid molecule for detecting at least one SNP given in the specification
CC comprising at least about 12 contiguous nucleotides, genotyping at least
CC one SNP position given in the specification in a sample, identifying an
CC individual having or at risk of developing a disorder and a kit
CC comprising at least one container containing the detection reagent.
CC Determining whether a trait is linked to one of the human chromosomes or
CC its sub-region comprises determining whether the trait is linked to one
CC or more SNPs using the detection reagents. Genotyping at least one SNP
CC position given in the specification in a sample comprises contacting the
CC sample with a detection reagent that differentiates between alternative
CC alleles at at least one SNP position given in the specification, and
CC determining which allele is present at the at least one SNP position.
CC Identifying an individual having or at risk of developing a disorder
CC comprises genotyping at least one SNP given in the specification in a
CC nucleic acid sample from the individual. The disorder is type II diabetes
CC (non-insulin dependent diabetes) or obesity. The detection reagent is
CC useful in identifying an individual having or at risk of developing a
CC disorder, particularly type II diabetes or obesity. This sequence
CC represents a human DNA polymorphic region used in the scope of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification but was obtained in electronic format from
CC USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 601 BP; 133 A; 150 C; 118 G; 200 T; 0 U; 0 Other;

Query Match 55.2%; Score 16; DB 14; Length 601;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 CAGGCATGAGCCAGCA 29
||||||||||||||||
Db 424 CAGGCATGAGCCAGCA 439

RESULT 25
ADB81754/c
ID ADB81754 standard; cDNA; 768 BP.
XX
AC ADB81754;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human cDNA sequence useful for the treatment of cancer (SeqID 66).
XX
KW human; prostate; cancer; cytostatic; gene therapy; vaccine;
KW immune response; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003050236-A2.
XX
PD 19-JUN-2003.
XX
PF 04-SEP-2002; 2002WO-US028214.
XX
PR 07-DEC-2001; 2001US-00012697.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI Crkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;
PI Garcia V, Jones LW, Stache-Crain B, Scott EM;
XX
DR WPI; 2003-513972/48.
XX
PT New polynucleotides derived from human prostate, useful for modulating
PT immune response to prevent or treat cancer.
XX
PS Claim 1; SEQ ID NO 66; 188pp; English.
XX
CC This invention relates to novel isolated polynucleotides of human origin,
CC particularly isolated from the human prostate. Specifically, it refers to
CC the diagnostics and therapeutics comprising these novel human
CC polynucleotides, and includes the derived probes, antisense
CC oligonucleotides and antibodies thereof. The identification of these
CC human prostate genes that can inhibit tumour growth is useful for
CC understanding the progression and nature of complex diseases such as
CC cancer, and hence they are important in the drug discovery process. The
CC present invention describes these polynucleotides and encoded
CC polypeptides as exhibiting cytostatic activity, and through gene therapy
CC and/ or vaccines they can be used to modulate the immune response for the
CC prevention or treatment of cancers, particularly of the prostate, but
CC also for breast, lung and colon cancer. This polynucleotide sequence is a
CC human cDNA sequence useful for the treatment of cancer, used in an

```
CC   exemplification of the invention. NOTE: These sequences are not given in
CC   the specification but are provided on the WIPO website.
XX
SQ   Sequence 768 BP; 215 A; 140 C; 153 G; 230 T; 0 U; 30 Other;

  Query Match             55.2%;  Score 16;  DB 9;  Length 768;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        543 CAGGCATGAGCCAGCA 528


RESULT 26
AAH03489
ID   AAH03489 standard; cDNA; 809 BP.
XX
AC   AAH03489;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA clone (5'-primer) SEQ ID NO:324.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   27-AUG-1999;   99JP-00300253.
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI   Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 1; SEQ ID NO 324; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesising 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 809 BP; 179 A; 182 C; 156 G; 289 T; 0 U; 3 Other;

  Query Match             55.2%;  Score 16;  DB 4;  Length 809;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 GCTCAGGCATGAGCCA 26
              ||||||||||||||||
Db         73 GCTCAGGCATGAGCCA 88


RESULT 27
ABQ88810/c
ID   ABQ88810 standard; cDNA; 817 BP.
XX
AC   ABQ88810;
XX
DT   27-SEP-2002  (first entry)
XX
DE   Human prostate expressed polynucleotide SEQ ID NO 66.
XX
KW   Human; prostate; cytostatic; tumour; cancer; vaccine; gene therapy; gene;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200255700-A2.
XX
PD   18-JUL-2002.
XX
PF   07-DEC-2001; 2001WO-US047349.
XX
PR   07-DEC-2000; 2000US-0254648P.
PR   13-MAR-2001; 2001US-0275688P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI   Crkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;
PI   Garcia V, Jones WL, Stache-Crain B, Scott EM;
XX
DR   WPI; 2002-557824/59.
XX
PT   New genes and gene products isolated from human prostate, useful for
PT   treating or diagnosing tumor or cancer (e.g. prostate cancer or breast
PT   cancer), or as vaccines for treating or preventing these diseases.
XX
PS   Claim 1; SEQ ID NO 66; 186pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any of
CC   1477 sequences or its fragment, degenerate variant, antisense or
CC   complement. The polynucleotides and gene products are useful for treating
CC   or diagnosing tumour or cancer (e.g. prostate cancer, breast cancer, lung
CC   cancer or medullary carcinoma) in a subject (e.g. cattle, dogs, cats,
CC   rabbits, horse or human). The polynucleotides and polypeptides are also
CC   useful as vaccines for treating or preventing these diseases. The
CC   polynucleotides are useful for gene therapy. The present sequence is that
CC   of one of a group of polynucleotides (ABQ88745-ABQ90015) disclosed
CC   electronically as sequences of the invention. However only 1271
CC   polynucleotide sequences are given, whereas 1477 polynucleotides and 91
CC   proteins are claimed. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequence
XX
SQ   Sequence 817 BP; 234 A; 151 C; 162 G; 232 T; 0 U; 38 Other;

  Query Match             55.2%;  Score 16;  DB 6;  Length 817;
  Best Local Similarity   100.0%;  Pred. No. 68;
```

```
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        562 CAGGCATGAGCCAGCA 547


RESULT 28
AAH07593/c
ID   AAH07593 standard; cDNA; 863 BP.
XX
AC   AAH07593;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA clone (5'-primer) SEQ ID NO:4428.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   27-AUG-1999;   99JP-00300253.
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 1; SEQ ID NO 4428; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesising 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 863 BP; 296 A; 135 C; 163 G; 261 T; 0 U; 8 Other;

  Query Match             55.2%;  Score 16;  DB 4;  Length 863;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        131 CAGGCATGAGCCAGCA 116


RESULT 29
AAK84716/c
ID   AAK84716 standard; DNA; 991 BP.
XX
AC   AAK84716;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:39528.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
```

PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 39528; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 991 BP; 346 A; 218 C; 181 G; 246 T; 0 U; 0 Other;

Query Match 55.2%; Score 16; DB 4; Length 991;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 CAGGCATGAGCCAGCA 29
||||||||||||||||
Db 744 CAGGCATGAGCCAGCA 729

RESULT 30
ABX15078/c
ID ABX15078 standard; cDNA; 1166 BP.
XX
AC ABX15078;
XX
DT 18-MAR-2003 (first entry)
XX
DE cDNA encoding novel polypeptide ribosomal protein S146.28.
XX
KW Gene; ss; polypeptide-ribosomal protein S416.28; tumour; haemopathy;
KW human immunodeficiency virus; HIV; infection; immunological disease;
KW inflammation.

```
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             419. .865
FT                   /*tag=  a
FT                   /product= "Polypeptide-ribosomal protein S416.28"
XX
PN   CN1358742-A.
XX
PD   17-JUL-2002.
XX
PF   13-DEC-2000; 2000CN-00127896.
XX
PR   13-DEC-2000; 2000CN-00127896.
XX
PA   (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2002-733531/80.
DR   P-PSDB; ABG72588.
XX
PT   Novel polypeptide-ribosomal S4 16.28 and polynucleotide for encoding the
PT   polypeptide.
XX
PS   Claim 6; Page 27-28 (disclosure); 34pp; Chinese.
XX
CC   The present invention discloses a novel polypeptide-ribosomal protein
CC   S416.28, the polynucleotide encoding this polypeptide and a method for
CC   producing this polypeptide by using DNA recombination technology. The
CC   invention also discloses the method for curing several diseases, such as
CC   malignant tumour, haemopathy, human immunodeficiency virus (HIV)
CC   infection, immunological disease and various inflammations by using the
CC   polypeptide. The invention also discloses an antagonist for resisting the
CC   polypeptide and its therapeutic action, and also discloses the
CC   application of polynucleotide coding this novel ribosomal protein
CC   S416.28. The present sequence represents the cDNA sequence encoding the
CC   ribosomal protein S416.28 protein of the invention
XX
SQ   Sequence 1166 BP; 337 A; 274 C; 326 G; 229 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 6;  Length 1166;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        252 CAGGCATGAGCCAGCA 237


RESULT 31
ADC87218/c
ID   ADC87218 standard; DNA; 1201 BP.
XX
AC   ADC87218;
XX
DT   01-JAN-2004  (first entry)
XX
DE   Human GPCR gene SEQ ID NO:1671.
XX
KW   ds; gene; human; GPCR;
KW   guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX
OS   Homo sapiens.
XX
PN   EP1270724-A2.
XX
PD   02-JAN-2003.
XX
PF   18-JUN-2002; 2002EP-00013517.
XX
PR   18-JUN-2001; 2001JP-00246789.
XX
PA   (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA   (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI   Suwa M,  Asai K,  Akiyama Y,  Aburatani H;
XX
DR   WPI; 2003-315783/31.
DR   P-PSDB; ADC87219.
XX
PT   New polynucleotide, useful for preparing a composition for treating a
PT   patient in need of increased or suppressed activity or expression of the
PT   guanosine triphosphate-binding protein coupled receptor.
XX
PS   Claim 1; SEQ ID NO 1671; 28pp; English.
XX
CC   The invention relates to a novel polynucleotide encoding a guanosine
CC   triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
CC   the invention may have a use in gene therapy. The polynucleotide and
CC   polypeptide are useful for preparing a composition for treating a patient
CC   in need of increased or suppressed activity or expression of the
CC   guanosine triphosphate-binding protein coupled receptor. The
CC   polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
CC   invention.
XX
SQ   Sequence 1201 BP; 141 A; 290 C; 341 G; 429 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 10;  Length 1201;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        996 CAGGCATGAGCCAGCA 981


RESULT 32
ADT48620/c
ID   ADT48620 standard; cDNA; 1386 BP.
XX
AC   ADT48620;
XX
DT   02-DEC-2004  (first entry)
XX
DE   Bacterial polynucleotide #23371.
XX
KW   Recombinant DNA construct; transformed plant; improved plant property;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW   pathogen tolerance; pest tolerance; plant disease resistance;
KW   cell cycle pathway modification; plant growth regulator;
KW   homologous recombination; seed oil yield; protein yield; carbohydrate;
KW   nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW   bacterial polynucleotide; gene; ss.
XX
OS   Bacteria.
XX
PN   US2003233675-A1.
XX
PD   18-DEC-2003.
XX
PF   20-FEB-2003; 2003US-00369493.
XX
PR   21-FEB-2002; 2002US-0360039P.
XX
PA   (CAOY/) CAO Y.
PA   (HINK/) HINKLE G J.
PA   (SLAT/) SLATER S C.
PA   (CHEN/) CHEN X.
PA   (GOLD/) GOLDMAN B S.
XX
PI   Cao Y,  Hinkle GJ,  Slater SC,  Chen X,  Goldman BS;
XX
DR   WPI; 2004-061375/06.
XX
```

```
PT   New recombinant DNA construct comprising a promoter positioned to provide
PT   for expression of a polynucleotide encoding a polypeptide from a
PT   microbial source, useful for producing plants with improved properties.
XX
PS   Claim 1; SEQ ID NO 47058; 122pp; English.
XX
CC   The invention relates to a recombinant DNA construct comprising a
CC   promoter functional in a plant cell, where the promoter is positioned to
CC   provide for expression of a polynucleotide encoding a polypeptide from a
CC   microbial source. The invention also relates to a transformed plant
CC   comprising the recombinant DNA construct and a method of producing a
CC   transformed plant having an improved property. The plant is a crop plant
CC   such as maize or soybean. The method of producing a transformed plant
CC   having an improved property comprises transforming a plant with the
CC   recombinant DNA construct and growing the transformed plant, where the
CC   polynucleotide or polypeptide is useful for improving plant properties.
CC   The recombinant DNA construct is useful for producing plants with
CC   improved plant properties, e.g. improved cold, heat or drought tolerance,
CC   tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC   increased resistance to plant disease, better growth rate by modification
CC   of the cell cycle pathway with plant growth regulators, increased rate of
CC   homologous recombination, modified seed oil or protein yield and/or
CC   content, improved yield by modification of carbohydrate, nitrogen or
CC   phosphorus use and/or uptake, by modification of photosynthesis or by
CC   providing improved plant growth and development under at least one stress
CC   condition, improved lignin production or improved galactomannan
CC   production. This sequence represents a bacterial polynucleotide used in
CC   the scope of the invention. Note: The sequence data for this patent did
CC   not form part of the printed specification but was obtained in electronic
CC   format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 1386 BP; 326 A; 336 C; 362 G; 362 T; 0 U; 0 Other;

  Query Match           55.2%;  Score 16;  DB 13;  Length 1386;
  Best Local Similarity  100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 GAGCTCAGGCATGAGC 24
              ||||||||||||||||
Db       1101 GAGCTCAGGCATGAGC 1086


RESULT 33
AAH15274
ID   AAH15274 standard; cDNA; 1610 BP.
XX
AC   AAH15274;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:13414.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   27-AUG-1999;   99JP-00300253.
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 8; SEQ ID NO 13414; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesising 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 1610 BP; 390 A; 379 C; 355 G; 486 T; 0 U; 0 Other;

  Query Match           55.2%;  Score 16;  DB 4;  Length 1610;
  Best Local Similarity  100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 GCTCAGGCATGAGCCA 26
              ||||||||||||||||
Db         73 GCTCAGGCATGAGCCA 88


RESULT 34
AAC69417/c
ID   AAC69417 standard; cDNA; 1638 BP.
XX
AC   AAC69417;
XX
DT   30-JAN-2001  (first entry)
XX
DE   Human secreted protein gene 19 SEQ ID NO:29.
XX
KW   Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW   antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KW   cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW   fungicide; ophthalmological; gene therapy; autoimmune disease; infection;
KW   hyperproliferative disorder; cardiovascular disorder; angiogenesis;
KW   cerebrovascular disorder; nervous system disorder; ocular disorder;
KW   wound healing; skin aging; food additive; preservative; ss.
XX
OS   Homo sapiens.
XX
PN   WO200058468-A2.
XX
PD   05-OCT-2000.
XX
PF   22-MAR-2000; 2000WO-US007526.
XX
PR   26-MAR-1999;   99US-0126600P.
PR   22-DEC-1999;   99US-0171550P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
```

```
XX
PI   Rosen CA,  Ruben SM,  Komatsoulis G;
XX
DR   WPI; 2000-611713/58.
DR   P-PSDB; AAB38137.
XX
PT   Nucleic acids encoding human secreted proteins, used to prevent, treat,
PT   ameliorate, or diagnose conditions such as autoimmune disorders, skin
PT   disorders and cancer.
XX
PS   Claim 1; Page 327; 374pp; English.
XX
CC   The polynucleotide sequences given in AAC69399 to AAC69445 encode the
CC   human secreted proteins given in AAB38119 to AAB38165. AAB38166 to
CC   AAB38201 represent human secreted polypeptide sequences and proteins
CC   homologous to them, which are given in the exemplification of the present
CC   invention. Human secreted proteins have activities based on the tissues
CC   and cells the genes are expressed in. Example of activities include:
CC   immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
CC   cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
CC   neuroprotective; antibacterial; virucide; fungicide; and
CC   ophthalmological. The polynucleotides and polypeptides can be are used to
CC   prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC   rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC   in diagnosing a pathological condition or susceptibility to a
CC   pathological condition. Disorders which are diagnosed or treated include
CC   autoimmune diseases, hyperproliferative disorders, cardiovascular
CC   disorders, cerebrovascular disorders, angiogenesis, nervous system
CC   disorders, infections caused by bacteria, viruses and fungi and ocular
CC   disorders. The polypeptides can also be used to aid wound healing and
CC   epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC   maintain organs before transplantation, for supporting cell culture of
CC   primary tissues, to regenerate tissues and in chemotaxis. The
CC   polypeptides can also be used as a food additive or preservative to
CC   increase or decrease storage capabilities. AAC69390 to AAC69398 and
CC   AAB38118 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 1638 BP; 496 A; 297 C; 331 G; 512 T; 0 U; 2 Other;

  Query Match             55.2%;  Score 16;  DB 3;  Length 1638;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db       1507 CAGGCATGAGCCAGCA 1492


RESULT 35
AAH14794
ID   AAH14794 standard; cDNA; 1911 BP.
XX
AC   AAH14794;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:12580.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   27-AUG-1999;   99JP-00300253.
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 8; SEQ ID NO 12580; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesising 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 1911 BP; 550 A; 368 C; 372 G; 621 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 4;  Length 1911;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        947 CAGGCATGAGCCAGCA 962


RESULT 36
ADC87088/c
ID   ADC87088 standard; DNA; 1989 BP.
XX
AC   ADC87088;
XX
DT   01-JAN-2004  (first entry)
XX
DE   Human GPCR gene SEQ ID NO:1541.
XX
KW   ds; gene; human; GPCR;
KW   guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX
OS   Homo sapiens.
XX
PN   EP1270724-A2.
XX
PD   02-JAN-2003.
XX
PF   18-JUN-2002; 2002EP-00013517.
XX
PR   18-JUN-2001; 2001JP-00246789.
XX
```

```
PA   (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA   (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI   Suwa M,  Asai K,  Akiyama Y,  Aburatani H;
XX
DR   WPI; 2003-315783/31.
DR   P-PSDB; ADC87089.
XX
PT   New polynucleotide, useful for preparing a composition for treating a
PT   patient in need of increased or suppressed activity or expression of the
PT   guanosine triphosphate-binding protein coupled receptor.
XX
PS   Claim 1; SEQ ID NO 1541; 28pp; English.
XX
CC   The invention relates to a novel polynucleotide encoding a guanosine
CC   triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
CC   the invention may have a use in gene therapy. The polynucleotide and
CC   polypeptide are useful for preparing a composition for treating a patient
CC   in need of increased or suppressed activity or expression of the
CC   guanosine triphosphate-binding protein coupled receptor. The
CC   polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
CC   invention.
XX
SQ   Sequence 1989 BP; 564 A; 413 C; 448 G; 564 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 10;  Length 1989;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db       1840 CAGGCATGAGCCAGCA 1825


RESULT 37
ACC57312
ID   ACC57312 standard; cDNA; 2103 BP.
XX
AC   ACC57312;
XX
DT   27-JUN-2003  (first entry)
XX
DE   Zinc finger protein 11.55 encoding cDNA # SEQ ID 1.
XX
KW   Zinc finger protein; 11.55; human immunodeficiency virus; HIV; cancer;
KW   gene; ss.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             1474. .1791
FT                   /*tag=  a
FT                   /product= "zinc finger protein 11.55"
XX
PN   CN1363594-A.
XX
PD   14-AUG-2002.
XX
PF   05-JAN-2001; 2001CN-00105078.
XX
PR   05-JAN-2001; 2001CN-00105078.
XX
PA   (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2003-000323/01.
DR   P-PSDB; ABP60166.
XX
PT   Polypeptide-zinc finger protein 11.55 and polynucleotide encoding it.
XX
PS   Claim 6; Page 25-26 (disclosure); 33pp; Chinese.
XX
CC   The invention relates to a novel zinc finger protein designated 11.55.
CC   Also disclosed are the polynucleotide encoding it, and a process for
CC   preparing the polypeptide using DNA recombination techniques. The
CC   application of the polypeptide is in treating diseases such as cancer and
CC   human immunodeficiency virus (HIV) infection, The current sequence
CC   represents the zinc finger protein 11.55 encoding cDNA
XX
SQ   Sequence 2103 BP; 535 A; 502 C; 580 G; 486 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 10;  Length 2103;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db         77 CAGGCATGAGCCAGCA 92


RESULT 38
ADQ64710/c
ID   ADQ64710 standard; cDNA; 2159 BP.
XX
AC   ADQ64710;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Novel human cDNA sequence #1871.
XX
KW   ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KW   cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW   neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW   cancer.
XX
OS   Homo sapiens.
XX
PN   EP1440981-A2.
XX
PD   28-JUL-2004.
XX
PF   21-JAN-2004; 2004EP-00001196.
XX
PR   21-JAN-2003; 2003JP-00102206.
PR   09-MAY-2003; 2003JP-00131392.
XX
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Isogai T,  Sugiyama T,  Otsuki T,  Wakamatsu A,  Sato H,  Ishii S;
PI   Yamamoto J,  Isono Y,  Nagai K,  Irie R;
XX
DR   WPI; 2004-535376/52.
DR   P-PSDB; ADQ66898.
XX
PT   Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT   Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS   Claim 1; SEQ ID NO 1871; 2449pp; English.
XX
CC   The invention relates to 2495 novel polynucleotides (I) and their encoded
CC   polypeptides, sequences hybridizing to these nucleotides, sequences
CC   encoding partial polypeptides and sequences having 70% or 90% identity to
CC   the nucleotide and protein sequences. The nucleotides and polypeptides
CC   are useful as diagnostic markers or therapeutic target for the diseases
CC   or morbid states. They are also useful for treating osteoporosis,
CC   neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC   dementia and various cancers. This sequence corresponds to a nucleotide
CC   sequence of the invention.
XX
SQ   Sequence 2159 BP; 598 A; 514 C; 553 G; 494 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 12;  Length 2159;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        1636 CAGGCATGAGCCAGCA 1621
```

RESULT 39
AAH14861/c
ID AAH14861 standard; cDNA; 2756 BP.
XX
AC AAH14861;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:12702.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 12702; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 2756 BP; 911 A; 438 C; 542 G; 865 T; 0 U; 0 Other;

```
  Query Match             55.2%;  Score 16;  DB 4;  Length 2756;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db         131 CAGGCATGAGCCAGCA 116
```

RESULT 40
ADR07944/c
ID ADR07944 standard; cDNA; 3015 BP.
XX
AC ADR07944;
XX
DT 04-NOV-2004 (first entry)
XX
DE Full length human cDNA useful for treating neurological disease Seq 1450.
XX
KW gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;
KW osteoporosis; neurological disease; Alzheimer's disease;
KW Parkinson's disease; dementia; short memory; cancer;
KW sense or motor function; emotional reaction; fear response; panic;
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW tranquiliser.
XX
OS Homo sapiens.
XX
PN EP1447413-A2.
XX
PD 18-AUG-2004.
XX
PF 12-FEB-2004; 2004EP-00003145.
XX
PR 14-FEB-2003; 2003JP-00102207.
PR 09-MAY-2003; 2003JP-00131452.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
PI Wakamatsu A, Ishii S, Nagai K, Irie R;
XX
DR WPI; 2004-583265/57.
DR P-PSDB; ADR09900.
XX
PT New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS Claim 1; SEQ ID NO 1450; 2686pp; English.
XX
CC This invention relates to novel, isolated full length human cDNA
CC molecules and the encoded proteins thereof. Specifically, it refers to
CC cDNA clones obtained by an oligo-capping method, where none of these
CC clones are identical to any known human mRNAs. The present invention
CC describes an immunoassay to identify agonists and antagonists, as well as
CC antibodies, antisense molecules and siRNAs that can all be used to bind
CC to and modulate expression of the cDNA molecules. As such, these
CC molecules are useful for diagnostic markers or therapeutic targets for
CC the various diseases or morbid states. In particular, they are useful in
CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC disease, Parkinson's disease, dementia, short memory and various cancers,
CC as well as for maintaining equilibrium of sense or motor function, and
CC for treating emotional reaction, fear response and panic. Accordingly,
CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
CC cytostatic and tranquiliser activities. This polynucleotide is a full
CC length human cDNA sequence of the invention. NOTE: This sequence is not
CC given in the sequence listing of the specification but can be obtained on
CC CD-ROM from the European Patent Office, Vienna Sub-office.
XX
SQ Sequence 3015 BP; 685 A; 874 C; 910 G; 546 T; 0 U; 0 Other;

```
  Query Match             55.2%;  Score 16;  DB 13;  Length 3015;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          14 CAGGCATGAGCCAGCA 29
```

```
                  ||||||||||||||||
Db       2760 CAGGCATGAGCCAGCA 2745


RESULT 41
ADB62850
ID   ADB62850 standard; cDNA; 3187 BP.
XX
AC   ADB62850;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human cDNA encoding clone OCBBF20117220.
XX
KW   Human; ss; gene; pharmaceutical; diagnostic; gene therapy;
KW   tissue regeneration; cell regeneration; membrane protein;
KW   signal transduction-related protein; transcription-related protein;
KW   osteoporosis; neurological disease; cancer; tumour.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             605. .937
FT                   /*tag= a
FT                   /product= "Clone OCBBF20117220 protein"
XX
PN   EP1308459-A2.
XX
PD   07-MAY-2003.
XX
PF   28-MAR-2002; 2002EP-00007401.
XX
PR   05-NOV-2001; 2001JP-00379298.
PR   25-JAN-2002; 2002US-00350978.
XX
PA   (HELI-) HELIX RES INST.
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Isogai T,  Sugiyama T,  Otsuki T,  Wakamatsu A,  Sato H,  Ishii S;
PI   Yamamoto J,  Isono Y,  Hio Y,  Otsuka K,  Nagai K,  Irie R,  Tamechika I;
PI   Seki N,  Yoshikawa T,  Otsuka M,  Nagahari K,  Masuho Y;
XX
DR   WPI; 2003-450961/43.
DR   P-PSDB; ADB64820.
XX
PT   New polynucleotides and polypeptides, useful for developing a diagnostic
PT   marker or medicines for regulation of their expression and activity, or
PT   as targets of gene therapy.
XX
PS   Claim 1; Page; 222pp; English.
XX
CC   The invention discloses a polynucleotide comprising a sequence selected
CC   from 1970 fully defined nucleotide sequences which encode novel
CC   polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC   or its partial peptide, an antibody binding to the polypeptide or peptide
CC   of the polynucleotide, immunologically assaying the polypeptide or
CC   peptide of the polynucleotide by contacting the polypeptide or peptide
CC   with the antibody of the encoded protein, and observing the binding
CC   between the two, a transformant carrying the polynucleotide in an
CC   expressible manner and an antisense polynucleotide. The oligonucleotide
CC   is useful as a primer for synthesising the polynucleotide, or as a probe
CC   for detecting the polynucleotide. The polynucleotides and encoded
CC   proteins are useful as pharmaceutical agents and many disease-related
CC   genes may be included in them, for developing a diagnostic marker or
CC   medicines for regulation of their expression and activity, or as targets
CC   of gene therapy. The genes are involved in tissue and/or cell
CC   regeneration. Membrane proteins, signal transduction-related proteins,
CC   transcription-related proteins, disease-related proteins and genes
CC   encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC   neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC   the activity or expression of the encoded protein to treat diseases. The
CC   sequence presented is a cDNA of the invention. Note: Some of the sequence
CC   data for this patent is not represented in the printed specification, but
CC   is based on sequence information supplied by the European Patent Office.
XX
SQ   Sequence 3187 BP; 768 A; 795 C; 795 G; 829 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 10;  Length 3187;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db       1740 CAGGCATGAGCCAGCA 1755


RESULT 42
AAK83192
ID   AAK83192 standard; DNA; 3608 BP.
XX
AC   AAK83192;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38004.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
```

PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 38004; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 3608 BP; 854 A; 890 C; 884 G; 980 T; 0 U; 0 Other;

Query Match 55.2%; Score 16; DB 4; Length 3608;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 CAGGCATGAGCCAGCA 29
||||||||||||||||
Db 1939 CAGGCATGAGCCAGCA 1954

RESULT 43
AAK74891
ID AAK74891 standard; DNA; 3608 BP.
XX
AC AAK74891;
XX
DT 07-NOV-2001 (first entry)
XX

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:29703.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.

PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 29703; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 3608 BP; 854 A; 890 C; 884 G; 980 T; 0 U; 0 Other;

Query Match 55.2%; Score 16; DB 4; Length 3608;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db       1939 CAGGCATGAGCCAGCA 1954
```

RESULT 44
AAK67271/c
ID AAK67271 standard; DNA; 3608 BP.
XX
AC AAK67271;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:22083.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.

```
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241786P.
PR   20-OCT-2000; 2000US-0241787P.
PR   20-OCT-2000; 2000US-0241808P.
PR   20-OCT-2000; 2000US-0241809P.
PR   20-OCT-2000; 2000US-0241826P.
PR   01-NOV-2000; 2000US-0244617P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
PR   08-NOV-2000; 2000US-0246523P.
PR   08-NOV-2000; 2000US-0246524P.
PR   08-NOV-2000; 2000US-0246525P.
PR   08-NOV-2000; 2000US-0246526P.
PR   08-NOV-2000; 2000US-0246527P.
PR   08-NOV-2000; 2000US-0246528P.
PR   08-NOV-2000; 2000US-0246532P.
PR   08-NOV-2000; 2000US-0246609P.
PR   08-NOV-2000; 2000US-0246610P.
PR   08-NOV-2000; 2000US-0246611P.
PR   08-NOV-2000; 2000US-0246613P.
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS   Disclosure; SEQ ID NO 22083; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting the
CC   nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention
XX
SQ   Sequence 3608 BP; 980 A; 884 C; 890 G; 854 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 4;  Length 3608;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        1670 CAGGCATGAGCCAGCA 1655


RESULT 45
AAF93891
ID   AAF93891 standard; cDNA; 3791 BP.
XX
AC   AAF93891;
XX
DT   23-MAY-2001  (first entry)
XX
DE   Human cDNA encoding a membrane or secretory protein clone PSEC0258.
XX
KW   Human; secretory protein; membrane protein; vaccine; gene therapy;
KW   rheumatoid arthritis; diabetes; ss.
XX
OS   Homo sapiens.
XX
PN   EP1067182-A2.
XX
PD   10-JAN-2001.
XX
PF   07-JUL-2000; 2000EP-00114090.
XX
PR   08-JUL-1999;   99JP-00194179.
PR   11-JAN-2000; 2000JP-00118775.
PR   02-MAY-2000; 2000JP-00183766.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Kawai Y,  Sugiyama T,  Hayashi K;
XX
DR   WPI; 2001-093989/11.
DR   P-PSDB; AAB88464.
XX
PT   Nucleic acids encoding secretory proteins/membrane proteins, useful in
PT   gene therapy or as candidate target molecules in drug development.
XX
PS   Claim 1; SEQ ID NO 295; 609pp + Sequence Listing; English.
XX
CC   This invention relates to nucleic acid sequences AAF93744 - AAF93916
CC   which encode human secretory or membrane proteins represented by AAB88317
CC   - AAB88419. Included in the invention are primers AAF93917 - AAF94295 and
CC   AAF62232 - AAF62235 which are used to isolate the cDNA sequences of the
CC   invention. The invention also includes methods for the production of
CC   antibodies directed against the proteins, and cDNA sequences, which can
CC   be used in vaccines. The polynucleotide sequences can be used in gene
CC   therapy. The polynucleotide sequences and the proteins they encode may be
CC   used in the prevention, treatment and diagnosis of diseases associated
CC   with inappropriate secretory protein/membrane protein expression. The
CC   nucleic acids and complementary sequences may also be used as DNA probes
CC   in diagnostic assays (e.g. polymerase chain reactions (PCR)) to detect
CC   and quantitate the presence of similar nucleic acid sequences in samples.
```

```
CC   They may also be used to study the expression and function of secretory
CC   proteins/membrane polypeptides and their role in metabolism. The
CC   polypeptides may be used as antigens in the production of antibodies
CC   against them and in assays to identify modulators (agonists and
CC   antagonists) of expression and activity. The antibodies and antagonists
CC   may also be used as therapeutic agents to down regulate expression and
CC   activity. The antibodies may also be used as diagnostic agents for
CC   detecting the presence of the polypeptides in samples (e.g. by enzyme
CC   linked immunosorbant assay (ELISA). Examples of diseases which may be
CC   treated include rheumatoid arthritis and diabetes
XX
SQ   Sequence 3791 BP; 894 A; 1055 C; 1079 G; 763 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 5;  Length 3791;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db       1802 CAGGCATGAGCCAGCA 1817


RESULT 46
ADY63290
ID   ADY63290 standard; cDNA; 3791 BP.
XX
AC   ADY63290;
XX
DT   02-JUN-2005  (first entry)
XX
DE   Human clone PSEC0258 cDNA, SEQ ID 295.
XX
KW   Gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   EP1514933-A1.
XX
PD   16-MAR-2005.
XX
PF   07-JUL-2000; 2004EP-00027228.
XX
PR   08-JUL-1999;   99JP-00194179.
PR   11-JAN-2000; 2000JP-00118775.
PR   02-MAY-2000; 2000JP-00183766.
PR   07-JUL-2000; 2000EP-00114090.
XX
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Kawai Y,  Sugiyama T,  Hayashi K;
XX
DR   WPI; 2005-203865/22.
DR   P-PSDB; ADY63291.
XX
PT   Novel isolated polynucleotide encoding human secretary proteins or
PT   membrane proteins, useful for examination and diagnosis of abnormality of
PT   human secretary proteins.
XX
PS   Disclosure; SEQ ID NO 295; 1240pp; English.
XX
CC   The present invention relates to novel human secretory proteins or
CC   membrane proteins, and their coding sequences. The present sequence is
CC   one such coding sequence. The coding sequences of the invention are
CC   useful for examination and diagnosis of abnormality of the human
CC   secretary proteins and in gene therapy methods. The coding sequences and
CC   proteins are useful as candidates for medicines or as target molecules
CC   for developing medicines. Antibodies against the proteins of the
CC   invention are useful for treating diseases that are associated with the
CC   proteins. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained from sequence information
CC   supplied by the European Patent Office.
XX
SQ   Sequence 3791 BP; 894 A; 1055 C; 1079 G; 763 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 14;  Length 3791;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db       1802 CAGGCATGAGCCAGCA 1817


RESULT 47
ABD07975
ID   ABD07975 standard; DNA; 3810 BP.
XX
AC   ABD07975;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Pseudomonas aeruginosa polynucleotide #6579.
XX
KW   Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KW   antibacterial.
XX
OS   Pseudomonas aeruginosa.
XX
PN   US6551795-B1.
XX
PD   22-APR-2003.
XX
PF   18-FEB-1999;   99US-00252991.
XX
PR   18-FEB-1998;   98US-0074788P.
PR   27-JUL-1998;   98US-0094190P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Rubenfield MJ,  Nolling J,  Deloughery C,  Bush D;
XX
DR   WPI; 2003-615309/58.
DR   P-PSDB; ABO74404.
XX
PT   Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT   useful as molecular targets for diagnostics, prophylaxis and treatment of
PT   pathological conditions resulting from bacterial infection.
XX
PS   Disclosure; SEQ ID NO 6579; 455pp; English.
XX
CC   The invention relates to Pseudomonas aeruginosa polypeptides and the
CC   polynucleotides encoding them. The sequences are useful in diagnosis and
CC   therapy of pathological conditions, as molecular targets for diagnostics,
CC   prophylaxis and treatment of pathological conditions resulting from a
CC   bacterial infection, for evaluating a compound, such as a polypeptide,
CC   for the ability to bind a P. aeruginosa nucleic acid, as components of
CC   effective antibacterial targets, as targets for antibacterial drugs,
CC   including anti-P. aeruginosa drugs, as templates for recombinant
CC   production of P. aeruginosa-derived peptides or polypeptides, as target
CC   components for diagnosis and/or treatment of P. aeruginosa-caused
CC   infection, and in detection of P. aeruginosa sequences or other sequences
CC   of Pseudomonas species using biochip technology. Sequences ABD01397-
CC   ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format from USPTO at
CC   seqdata.uspto.gov/sequence.html
XX
SQ   Sequence 3810 BP; 599 A; 1326 C; 1328 G; 557 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 11;  Length 3810;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db         68 CAGGCATGAGCCAGCA 83
```

```
RESULT 48
ADI31783/c
ID   ADI31783 standard; cDNA; 4428 BP.
XX
AC   ADI31783;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Human cDNA #1109.
XX
KW   Human; gene; ss; immunological response; immunopathological condition;
KW   Crohn's disease; asthma; ulcerative colitis; hypereosinophilia;
KW   irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;
KW   acute monocytic leukaemia; antiinflammatory; antiasthmatic; antiulcer;
KW   osteopathic; antiarthritic; antirheumatic; cytostatic.
XX
OS   Homo sapiens.
XX
PN   US6607879-B1.
XX
PD   19-AUG-2003.
XX
PF   09-FEB-1998;   98US-00023655.
XX
PR   09-FEB-1998;   98US-00023655.
XX
PA   (INCY-) INCYTE CORP.
XX
PI   Cocks BG,  Stuart SG,  Seilhamer JJ;
XX
DR   WPI; 2003-895307/82.
XX
PT   A composition comprising a plurality of cDNAs, useful for detecting
PT   altered expression of genes in an immunological response or for
PT   diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma
PT   or osteoarthritis.
XX
PS   Claim 1; SEQ ID NO 1109; 50pp; English.
XX
CC   The invention relates to a composition comprising a plurality of cDNAs
CC   for detecting the altered expression of genes in an immunological
CC   response. The invention also relates to a method of diagnosing or
CC   monitoring the treatment of an immunopathological condition in a sample,
CC   comprising obtaining nucleic acids from a sample, contacting the nucleic
CC   acids of the sample with an array comprising the plurality of cDNAs under
CC   conditions to form one or more hybridisation complexes, detecting the
CC   hybridisation complexes and comparing the levels of the detected
CC   hybridisation complexes with the level of hybridisation complexes
CC   detected in a non-diseased sample, where an altered level of the detected
CC   hybridisation complexes correlates with the presence of an
CC   immunopathological condition. Also disclosed are an expression profile
CC   comprising a microarray and a plurality of detectable complexes and a
CC   method for identifying a plurality of polynucleotide probes. The cDNAs
CC   are useful as hybridisable array elements in a microarray for monitoring
CC   the expression of target polynucleotides. The microarray can be used in
CC   the diagnosis of an immunopathology, such as Crohn's disease, asthma,
CC   ulcerative colitis, hypereosinophilia, irritable bowel syndrome,
CC   osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in
CC   identifying agents for the treatment of the diseases. The microarray may
CC   also be used in drug discovery and development, toxicological and
CC   carcinogenicity studies, forensics or pharmacogenomics. The composition
CC   may also be used in purification of a subpopulation of mRNAs, cDNAs or
CC   genomic fragments. This sequence represents a human cDNA of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification but was obtained in electronic format directly
CC   from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 4428 BP; 951 A; 1266 C; 1251 G; 960 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 11;  Length 4428;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db       3728 CAGGCATGAGCCAGCA 3713


RESULT 49
ADS83850/c
ID   ADS83850 standard; cDNA; 4428 BP.
XX
AC   ADS83850;
XX
DT   11-AUG-2005  (first entry)
XX
DE   Human lymph node cDNA #1109.
XX
KW   ss; gene; human; immunological response; blood cell; cancer;
KW   immunopathological; AIDS; allergy; anaemia; asthma; atherosclerosis;
KW   bronchitis; ulcerative colitis; diabetes; multiple sclerosis;
KW   osteoporosis; pancreatitis; infection; arthritis; lymph node.
XX
OS   Homo sapiens.
XX
PN   US2004077003-A1.
XX
PD   22-APR-2004.
XX
PF   14-AUG-2003; 2003US-00641643.
XX
PR   09-FEB-1998;   98US-00023655.
XX
PA   (INCY-) INCYTE CORP.
XX
PI   Cocks BG,  Stuart SG,  Seilhamer JJ;
XX
DR   WPI; 2004-387937/36.
XX
PT   New compositions having a number of first, second and third
PT   polynucleotide probes, useful in research and diagnostic applications in
PT   cancer and immunological conditions e.g. AIDS, diabetes, osteoporosis and
PT   infections.
XX
PS   Claim 15; SEQ ID NO 1109; 16pp; English.
XX
CC   The invention relates to polynucleotides which are used as probes to
CC   detect genes differentially expressed in an immunological response,
CC   abundantly expressed in an immunological response and/or coding for a
CC   polypeptide known to regulate blood cell biology. The polynucleotides are
CC   useful in research and diagnostic applications particularly in cancer and
CC   immunopathological conditions, such as AIDS, allergies, anaemia, asthma,
CC   atherosclerosis, bronchitis, ulcerative colitis, diabetes, multiple
CC   sclerosis, osteoporosis, pancreatitis, infections and arthritis. The
CC   present sequence represents a human lymph node cDNA used to detect blood
CC   cell and immunological response gene expression. Note: The present
CC   sequence does not appear in the printed specification but was obtained in
CC   electronic format from the USPTO web site
CC   (seqdata.uspto.gov/sequence.html?DocID=20040077003).
XX
SQ   Sequence 4428 BP; 951 A; 1266 C; 1251 G; 960 T; 0 U; 0 Other;

  Query Match             55.2%;  Score 16;  DB 13;  Length 4428;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db       3728 CAGGCATGAGCCAGCA 3713


RESULT 50
ADE71240/c
ID   ADE71240 standard; DNA; 4891 BP.
```

XX
AC ADE71240;
XX
DT 29-JAN-2004 (first entry)
XX
DE Novel human protein coding sequence #56.
XX
KW human; novel protein; drug; gene; ds.
XX
OS Homo sapiens.
XX
PN JP2002345493-A.
XX
PD 03-DEC-2002.
XX
PF 29-MAR-2001; 2002JP-00049046.
XX
PR 29-MAR-2001; 2001JP-00095524.
XX
PA (KAZU-) ZH KAZUSA DNA KENKYUSHO.
XX
DR WPI; 2003-460885/44.
DR P-PSDB; ADE71302.
XX
PT A gene and a protein encoded by it, used in drugs.
XX
PS Claim 1; SEQ ID NO 57; 257pp; Japanese.
XX
CC The invention comprises the amino acid and coding sequences of novel
CC human proteins. The DNA and protein sequences of the invention are used
CC in drugs. The present DNA sequence encodes a novel human protein of the
CC invention.
XX
SQ Sequence 4891 BP; 1582 A; 819 C; 889 G; 1601 T; 0 U; 0 Other;

Query Match 55.2%; Score 16; DB 10; Length 4891;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 CAGGCATGAGCCAGCA 29
||||||||||||||||
Db 2673 CAGGCATGAGCCAGCA 2658

Search completed: April 12, 2006, 13:25:31
Job time : 335 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: April 12, 2006, 13:25:43 ; Search time 1829 Seconds
(without alignments)
901.290 Million cell updates/sec

Title: SEQ1-4023-4051-4037A
Perfect score: 29
Sequence: 1 cctctctggagctcaggcatgagccagca 29

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 5883141 seqs, 28421725653 residues

Word size : 15

Total number of hits satisfying chosen parameters: 4607

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	19	65.5	2145	6	AX300008	AX300008 Sequence
c	2	19	65.5	170301	14	AC069487	AC069487 Homo sapi
	3	19	65.5	176947	8	HS170A21	Z82189 Human DNA s
c	4	19	65.5	238521	14	AC078932	AC078932 Mus muscu
	5	19	65.5	247508	9	AC153817	AC153817 Mus muscu
c	6	18	62.1	2300	5	AF181894	AF181894 Taricha g
	7	18	62.1	34713	14	BX539329	BX539329 Homo sapi
	8	18	62.1	47139	8	AL450465	AL450465 Human DNA
	9	18	62.1	126329	9	AL928544	AL928544 Mouse DNA
c	10	18	62.1	161289	5	BX649367	BX649367 Zebrafish
c	11	18	62.1	164201	8	AL157935	AL157935 Human DNA
	12	18	62.1	170975	5	BX677664	BX677664 Zebrafish
	13	18	62.1	174243	14	AC155304	AC155304 Mus muscu
	14	18	62.1	182271	14	AC080169	AC080169 Mus muscu
c	15	18	62.1	200288	14	AC119725	AC119725 Homo sapi
	16	18	62.1	207669	5	AL954146	AL954146 Zebrafish
	17	18	62.1	221370	14	AC126205	AC126205 Rattus no
	18	18	62.1	248701	14	AC107010	AC107010 Rattus no
	19	17	58.6	650	10	BV323421	BV323421 S236P6539
	20	17	58.6	1006	8	F272846S27	AF273247 Homo sapi
c	21	17	58.6	68409	14	AC105975	AC105975 Mus muscu
c	22	17	58.6	72300	14	AC116403	AC116403 Mus muscu
c	23	17	58.6	79564	8	BX649443	BX649443 Human DNA
c	24	17	58.6	80914	8	AC007999	AC007999 Homo sapi
	25	17	58.6	100080	8	HSDJ999L4	AL132821 Human DNA
	26	17	58.6	102313	8	AL603749	AL603749 Human DNA
	27	17	58.6	103979	5	CR352288	CR352288 Zebrafish
	28	17	58.6	110000	14	AC091528_1	Continuation (2 of
c	29	17	58.6	120574	8	AC104812	AC104812 Homo sapi
c	30	17	58.6	121857	5	CR384099	CR384099 Zebrafish
	31	17	58.6	126083	5	CR339049	CR339049 Zebrafish
c	32	17	58.6	141273	8	AL611933	AL611933 Human DNA
c	33	17	58.6	147431	9	AL591936	AL591936 Mouse DNA
c	34	17	58.6	149817	14	AC069440	AC069440 Homo sapi
c	35	17	58.6	152405	8	AF186190	AF186190 Homo sapi
	36	17	58.6	152884	8	AC138625	AC138625 Homo sapi
c	37	17	58.6	153012	8	CR854849	CR854849 Human DNA
	38	17	58.6	154279	8	AC112229	AC112229 Homo sapi
c	39	17	58.6	154729	14	AC138879	AC138879 Homo sapi
	40	17	58.6	154732	14	AC023236	AC023236 Homo sapi
c	41	17	58.6	156826	14	AC141284	AC141284 Homo sapi
c	42	17	58.6	157289	8	AC069421	AC069421 Homo sapi
	43	17	58.6	157462	8	AC130459	AC130459 Homo sapi
c	44	17	58.6	157838	8	HUAC004626	AC004626 Homo sapi
	45	17	58.6	158420	8	AC137788	AC137788 Homo sapi
	46	17	58.6	162025	14	AC026533	AC026533 Homo sapi
c	47	17	58.6	162585	9	AC164010	AC164010 Mus muscu
	48	17	58.6	164068	9	AL808118	AL808118 Mouse DNA
	49	17	58.6	166258	14	AL451066	AL451066 Homo sapi
	50	17	58.6	166418	14	AC020990	AC020990 Homo sapi
	51	17	58.6	167457	14	AC037478	AC037478 Homo sapi
c	52	17	58.6	168119	9	AC118935	AC118935 Mus muscu
	53	17	58.6	168635	14	AC130401	AC130401 Homo sapi
c	54	17	58.6	169013	8	AC034193	AC034193 Homo sapi
c	55	17	58.6	172879	14	AC068487	AC068487 Homo sapi
	56	17	58.6	175047	8	AC013280	AC013280 Homo sapi
	57	17	58.6	175099	8	AC109135	AC109135 Homo sapi
c	58	17	58.6	176981	8	AC136634	AC136634 Homo sapi
c	59	17	58.6	177885	14	AC128258	AC128258 Rattus no
	60	17	58.6	178080	14	AC011919	AC011919 Homo sapi
	61	17	58.6	178230	8	AC018828	AC018828 Homo sapi
c	62	17	58.6	180615	14	AC025549	AC025549 Homo sapi
c	63	17	58.6	182600	14	AC023088	AC023088 Homo sapi
c	64	17	58.6	183434	8	AC146162	AC146162 Pan trogl
c	65	17	58.6	184591	8	AL365443	AL365443 Human DNA
	66	17	58.6	187077	8	AC145939	AC145939 Pan trogl
	67	17	58.6	188631	8	AC022383	AC022383 Homo sapi
	68	17	58.6	188719	8	AC048330	AC048330 Homo sapi
	69	17	58.6	189255	14	AC068520	AC068520 Homo sapi
	70	17	58.6	190562	9	AC159307	AC159307 Mus muscu
	71	17	58.6	193214	8	AC069461	AC069461 Homo sapi
	72	17	58.6	197061	9	AL606914	AL606914 Mouse DNA
	73	17	58.6	198014	14	AC130026	AC130026 Rattus no
c	74	17	58.6	198546	14	AC069337	AC069337 Homo sapi
	75	17	58.6	201753	14	CR751225	CR751225 Danio rer
c	76	17	58.6	203685	9	AC122339	AC122339 Mus muscu
c	77	17	58.6	205928	8	CNS07EFU	AL583722 Human chr
	78	17	58.6	206457	8	AC013268	AC013268 Homo sapi
c	79	17	58.6	206578	14	AC026529	AC026529 Homo sapi
c	80	17	58.6	207708	8	AC018808	AC018808 Homo sapi
	81	17	58.6	209987	14	AC146130	AC146130 Pan trogl
	82	17	58.6	210252	14	AC160248	AC160248 Pan trogl
	83	17	58.6	212160	14	CR381672	CR381672 Danio rer
c	84	17	58.6	213575	9	AC154438	AC154438 Mus muscu
c	85	17	58.6	215402	14	AC151859	AC151859 Bos tauru
	86	17	58.6	215819	8	AC140490	AC140490 Homo sapi
c	87	17	58.6	216368	14	CT025592	CT025592 Mus muscu
c	88	17	58.6	216689	14	CR762382	CR762382 Homo sapi
c	89	17	58.6	217454	14	AL592070	AL592070 Mus muscu
	90	17	58.6	218698	8	AC007126	AC007126 Homo sapi
	91	17	58.6	223199	14	AC156405	AC156405 Bos tauru

c 92	17	58.6	224138	9	AC109243	AC109243 Mus muscu
93	17	58.6	225435	14	AC130505	AC130505 Rattus no
c 94	17	58.6	227230	8	AC140479	AC140479 Homo sapi
c 95	17	58.6	230353	14	AC130642	AC130642 Rattus no
96	17	58.6	241640	14	AC142065	AC142065 Rattus no
97	17	58.6	243356	14	AC156651	AC156651 Bos tauru
c 98	17	58.6	246883	14	AC126110	AC126110 Rattus no
c 99	17	58.6	248772	14	AC146019	AC146019 Pan trogl
100	17	58.6	253750	14	AC105781	AC105781 Rattus no
c 101	17	58.6	259230	14	AC134069	AC134069 Rattus no
102	17	58.6	259682	14	AC111230	AC111230 Rattus no
103	16	55.2	113	10	G33019	G33019 A009Z25 Hum
104	16	55.2	362	10	BV090159	BV090159 RPAMMSEQ0
105	16	55.2	362	10	BV097851	BV097851 RPAMMSEQ0
106	16	55.2	362	10	BV155075	BV155075 RPAMMSEQ0
c 107	16	55.2	391	6	CQ524503	CQ524503 Sequence
108	16	55.2	410	6	CQ526820	CQ526820 Sequence
c 109	16	55.2	420	6	BD113429	BD113429 EST and e
c 110	16	55.2	420	6	AR417876	AR417876 Sequence
c 111	16	55.2	420	6	AX978570	AX978570 Sequence
112	16	55.2	426	6	CQ486923	CQ486923 Sequence
c 113	16	55.2	452	6	CQ439437	CQ439437 Sequence
c 114	16	55.2	459	6	CQ713371	CQ713371 Sequence
115	16	55.2	463	6	CQ052001	CQ052001 Sequence
116	16	55.2	463	6	CQ067044	CQ067044 Sequence
117	16	55.2	463	6	CQ094098	CQ094098 Sequence
118	16	55.2	463	6	CQ132880	CQ132880 Sequence
119	16	55.2	463	6	CQ171467	CQ171467 Sequence
120	16	55.2	463	6	CQ200581	CQ200581 Sequence
121	16	55.2	463	6	CQ216095	CQ216095 Sequence
122	16	55.2	463	6	CQ254656	CQ254656 Sequence
123	16	55.2	463	6	CQ291740	CQ291740 Sequence
124	16	55.2	463	6	CQ328800	CQ328800 Sequence
c 125	16	55.2	464	10	AB140601	AB140601 Homo sapi
c 126	16	55.2	495	6	CQ518551	CQ518551 Sequence
127	16	55.2	547	10	BV546313	BV546313 qod86g07.
128	16	55.2	549	10	BV505738	BV505738 qri07g10.
129	16	55.2	560	10	BV209888	BV209888 DNM1_4192
130	16	55.2	564	6	CQ516733	CQ516733 Sequence
c 131	16	55.2	573	6	AX388772	AX388772 Sequence
c 132	16	55.2	638	10	BV548466	BV548466 slz73g08.
c 133	16	55.2	638	10	BV551965	BV551965 S221P6166
134	16	55.2	640	8	AY559744	AY559744 Homo sapi
c 135	16	55.2	669	10	BV546396	BV546396 qoq07f05.
c 136	16	55.2	684	10	BV540234	BV540234 G591P6420
137	16	55.2	693	10	BV542689	BV542689 G591P5728
138	16	55.2	693	10	BV576969	BV576969 G591P6084
139	16	55.2	703	10	BV570344	BV570344 G591P6689
c 140	16	55.2	703	10	BV497342	BV497342 S217P6093
141	16	55.2	711	10	BV551698	BV551698 S221P6675
c 142	16	55.2	712	10	BV590389	BV590389 G591P6417
c 143	16	55.2	722	10	BV534138	BV534138 G591P6222
144	16	55.2	738	10	BV632466	BV632466 S215P6033
145	16	55.2	740	10	BV672063	BV672063 S217P6116
c 146	16	55.2	743	8	HSA331729	AJ331729 Homo sapi
c 147	16	55.2	765	10	BV596721	BV596721 S216P6182
c 148	16	55.2	781	10	BV667818	BV667818 S217P6065
c 149	16	55.2	782	10	BV548891	BV548891 soa56g04.
150	16	55.2	809	6	BD145481	BD145481 Primer fo
151	16	55.2	809	6	AX865419	AX865419 Sequence
c 152	16	55.2	817	6	AX540819	AX540819 Sequence
153	16	55.2	826	9	RNMRP	X96394 R.norvegicu
c 154	16	55.2	827	10	BV502442	BV502442 qji20c03.
155	16	55.2	828	10	BV549678	BV549678 S215P6117
c 156	16	55.2	834	10	BV544639	BV544639 roj31e01.
c 157	16	55.2	863	6	BD149585	BD149585 Primer fo
c 158	16	55.2	863	6	AX869523	AX869523 Sequence
159	16	55.2	867	10	BV647792	BV647792 S215P6139
c 160	16	55.2	868	10	BV544396	BV544396 rny72h07.
c 161	16	55.2	873	10	BV570432	BV570432 G591P6572
c 162	16	55.2	1201	6	AX647479	AX647479 Sequence
163	16	55.2	1383	6	CQ738403	CQ738403 Sequence
c 164	16	55.2	1479	8	AK000831	AK000831 Homo sapi
c 165	16	55.2	1500	5	CR389265	CR389265 Gallus ga
166	16	55.2	1610	6	BD157266	BD157266 Primer fo
167	16	55.2	1610	6	AX878509	AX878509 Sequence
168	16	55.2	1610	8	AK021573	AK021573 Homo sapi
c 169	16	55.2	1638	6	BD275470	BD275470 47 Human
170	16	55.2	1642	8	BC043215	BC043215 Homo sapi
171	16	55.2	1911	6	BD156786	BD156786 Primer fo
172	16	55.2	1911	6	AX877675	AX877675 Sequence
173	16	55.2	1911	8	AK027771	AK027771 Homo sapi
174	16	55.2	1947	9	AF213389	AF213389 Mus muscu
c 175	16	55.2	1989	6	AX647349	AX647349 Sequence
c 176	16	55.2	2083	8	HSTHSMT6	AF019368 Human thi
c 177	16	55.2	2159	6	CQ843224	CQ843224 Sequence
c 178	16	55.2	2159	8	AK126169	AK126169 Homo sapi
c 179	16	55.2	2467	8	HSU93721	U93721 Homo sapien
c 180	16	55.2	2756	6	BD156853	BD156853 Primer fo
c 181	16	55.2	2756	6	AX877797	AX877797 Sequence
c 182	16	55.2	2756	8	AK001906	AK001906 Homo sapi
183	16	55.2	2969	9	RNO277881	AJ277881 Rattus no
c 184	16	55.2	3015	6	CQ850981	CQ850981 Sequence
c 185	16	55.2	3015	8	AK128829	AK128829 Homo sapi
186	16	55.2	3187	6	AX747479	AX747479 Sequence
187	16	55.2	3187	8	AK092318	AK092318 Homo sapi
c 188	16	55.2	3257	8	BC007730	BC007730 Homo sapi
c 189	16	55.2	3504	8	HSA012453	AJ012453 Homo sapi
190	16	55.2	3791	6	BD123639	BD123639 Secretory
191	16	55.2	3791	6	CS051409	CS051409 Sequence
192	16	55.2	3791	6	AX136373	AX136373 Sequence
193	16	55.2	3791	8	AK075558	AK075558 Homo sapi
c 194	16	55.2	4428	6	AR380564	AR380564 Sequence
c 195	16	55.2	4428	8	HUMHLL4G	M57678 Human 14 kD
196	16	55.2	4854	9	AY174893	AY174893 Rattus no
c 197	16	55.2	4891	6	BD185213	BD185213 Novel gen
198	16	55.2	4954	9	AY174892	AY174892 Rattus no
199	16	55.2	4981	9	AF487549	AF487549 Rattus no
c 200	16	55.2	5504	5	BC073441	BC073441 Xenopus l
201	16	55.2	5889	6	AR070200	AR070200 Sequence
202	16	55.2	5889	6	AR093441	AR093441 Sequence
203	16	55.2	5959	9	AK129095	AK129095 Mus muscu
c 204	16	55.2	7411	8	AL513219	AL513219 Human DNA
205	16	55.2	10372	1	AE003953	AE003953 Xylella f
c 206	16	55.2	10493	1	AE007308	AE007308 Sinorhizo
c 207	16	55.2	13295	14	AC009029	AC009029 Homo sapi
208	16	55.2	14664	6	A51135	A51135 Sequence 4
209	16	55.2	14664	6	A76960	A76960 Sequence 4
210	16	55.2	14664	6	AR632336	AR632336 Sequence
211	16	55.2	15019	8	AL591667	AL591667 Human DNA
c 212	16	55.2	17712	1	AE004844	AE004844 Pseudomon
213	16	55.2	21971	14	AC016141	AC016141 Homo sapi
c 214	16	55.2	27724	8	BX284678	BX284678 Human DNA
c 215	16	55.2	27870	8	AB045146	AB045146 Homo sapi
c 216	16	55.2	28230	8	HSCG1160	Z46936 Human DNA s
217	16	55.2	29067	8	AL645480	AL645480 Human DNA
218	16	55.2	30180	8	AY095373	AY095373 Homo sapi
219	16	55.2	30436	8	AC093016	AC093016 Homo sapi
220	16	55.2	31466	8	AY038934	AY038934 Homo sapi
221	16	55.2	32246	8	HSG0453	Z83001 Human DNA s
c 222	16	55.2	32367	6	AX600639	AX600639 Sequence
223	16	55.2	32706	8	AF191070	AF191070 Homo sapi
224	16	55.2	32926	14	AL354742	AL354742 Homo sapi
225	16	55.2	34680	8	AC023159	AC023159 Homo sapi
226	16	55.2	34744	8	AL589842	AL589842 Human DNA
227	16	55.2	35143	8	AP000229	AP000229 Homo sapi
228	16	55.2	36219	8	AC020946	AC020946 Homo sapi
c 229	16	55.2	36252	8	AP001056	AP001056 Homo sapi
230	16	55.2	36448	1	AB005554	AB005554 Bacillus
c 231	16	55.2	37027	8	HS506G2B	Z82976 Human DNA s
c 232	16	55.2	37389	8	AL355149	AL355149 Human DNA
c 233	16	55.2	37476	8	HS316G12	AL031709 Human DNA
c 234	16	55.2	37522	8	AC011529	AC011529 Homo sapi
235	16	55.2	37738	8	AC124856	AC124856 Homo sapi
c 236	16	55.2	38096	8	BX901949	BX901949 Human DNA
c 237	16	55.2	38580	8	AL590811	AL590811 Human DNA

c	238	16	55.2	38727	8	AC006953 AC006953 Homo sapi
c	239	16	55.2	38849	8	AC005346 AC005346 Homo sapi
c	240	16	55.2	39374	8	AC005256 AC005256 Homo sapi
	241	16	55.2	39577	8	AC008975 AC008975 Homo sapi
	242	16	55.2	40451	8	U73644 U73644 Human Chrom
c	243	16	55.2	40592	8	HSL191F1 Z68756 Human DNA s
	244	16	55.2	41335	8	CR942175 CR942175 Human DNA
	245	16	55.2	41589	8	AC020955 AC020955 Homo sapi
c	246	16	55.2	42304	8	AL355881 AL355881 Human DNA
	247	16	55.2	42569	8	AC005306 AC005306 Homo sapi
	248	16	55.2	42571	6	AR235865 AR235865 Sequence
	249	16	55.2	42619	8	AC005615 AC005615 Homo sapi
c	250	16	55.2	43275	8	AC125232 AC125232 Homo sapi
	251	16	55.2	44380	8	AC074193 AC074193 Homo sapi
c	252	16	55.2	45695	14	AC087610 AC087610 Homo sapi
c	253	16	55.2	46625	6	CQ870416 CQ870416 Sequence
c	254	16	55.2	47003	8	AC005335 AC005335 Homo sapi
c	255	16	55.2	48106	8	AC018476 AC018476 Homo sapi
	256	16	55.2	48478	14	AC026687 AC026687 Homo sapi
	257	16	55.2	48606	14	AC025671 AC025671 Homo sapi
	258	16	55.2	48632	14	BX890606 BX890606 Danio rer
	259	16	55.2	49525	8	AY165178 AY165178 Homo sapi
c	260	16	55.2	49816	8	AC093149 AC093149 Homo sapi
c	261	16	55.2	49961	8	DQ062746 DQ062746 Homo sapi
c	262	16	55.2	49976	14	AC025670 AC025670 Homo sapi
c	263	16	55.2	50768	8	AF095901 AF095901 Homo sapi
	264	16	55.2	51000	8	AL158171 AL158171 Human DNA
	265	16	55.2	51090	14	AL390837 AL390837 Homo sapi
c	266	16	55.2	51815	14	AC068454 AC068454 Homo sapi
c	267	16	55.2	52479	14	AC074125 AC074125 Homo sapi
	268	16	55.2	52840	14	AC112773_3 Continuation (4 of
	269	16	55.2	53196	8	AC108047 AC108047 Homo sapi
c	270	16	55.2	53823	14	AC091642 AC091642 Homo sapi
c	271	16	55.2	54154	8	AL583845 AL583845 Human DNA
	272	16	55.2	54207	14	AP002024 AP002024 Homo sapi
c	273	16	55.2	55672	14	AC103991 AC103991 Homo sapi
c	274	16	55.2	55957	14	AC110155 AC110155 Mus muscu
	275	16	55.2	56083	8	AC108488 AC108488 Homo sapi
	276	16	55.2	56652	14	AC027427 AC027427 Homo sapi
	277	16	55.2	56686	8	AY902237 AY902237 Homo sapi
c	278	16	55.2	56999	14	AC008776 AC008776 Homo sapi
c	279	16	55.2	57662	14	AC107969 AC107969 Homo sapi
	280	16	55.2	58039	5	CR759821 CR759821 Zebrafish
	281	16	55.2	58272	8	AC121338 AC121338 Homo sapi
	282	16	55.2	58749	14	AC139006 AC139006 Homo sapi
	283	16	55.2	59564	14	AC129481 AC129481 Homo sapi
c	284	16	55.2	59836	8	AL929592 AL929592 Human DNA
	285	16	55.2	60197	14	AC127510 AC127510 Homo sapi
c	286	16	55.2	60358	14	U82668 U82668 Homo sapien
c	287	16	55.2	60637	14	AC011149 AC011149 Homo sapi
	288	16	55.2	60943	14	AC135620 AC135620 Homo sapi
c	289	16	55.2	61101	14	AC083975 AC083975 Homo sapi
	290	16	55.2	62124	6	CQ869935 CQ869935 Sequence
	291	16	55.2	62196	14	AC025830 AC025830 Homo sapi
	292	16	55.2	63187	8	AL355477 AL355477 Human DNA
	293	16	55.2	63441	8	AL356502 AL356502 Human DNA
c	294	16	55.2	63584	14	AC121321 AC121321 Homo sapi
	295	16	55.2	63939	8	AL359200 AL359200 Human DNA
c	296	16	55.2	63973	14	AC130351 AC130351 Homo sapi
c	297	16	55.2	64359	8	AC005233 AC005233 Homo sapi
c	298	16	55.2	65379	8	AP000404 AP000404 Homo sapi
	299	16	55.2	65461	14	AC131292 AC131292 Homo sapi
c	300	16	55.2	65461	14	AC131292 AC131292 Homo sapi
	301	16	55.2	66086	14	AC012249 AC012249 Homo sapi
c	302	16	55.2	66325	14	AC016425 AC016425 Homo sapi
	303	16	55.2	66741	8	HS460J8 AL031662 Human DNA
	304	16	55.2	67335	8	HS436B3 AL449212 Homo sapi
c	305	16	55.2	67407	14	AC087519 AC087519 Homo sapi
	306	16	55.2	67468	14	AC139148 AC139148 Homo sapi
	307	16	55.2	67674	14	AC090641 AC090641 Homo sapi
c	308	16	55.2	68299	8	AC091685 AC091685 Homo sapi
c	309	16	55.2	68378	14	AC113131 AC113131 Homo sapi
	310	16	55.2	68459	14	AC137964 AC137964 Homo sapi
	311	16	55.2	69085	8	AC009979 AC009979 Homo sapi
	312	16	55.2	69666	14	AC090144 AC090144 Homo sapi
c	313	16	55.2	69702	14	AC027743 AC027743 Homo sapi
	314	16	55.2	71819	8	AC007536 AC007536 Homo sapi
c	315	16	55.2	71864	14	AC103852 AC103852 Homo sapi
c	316	16	55.2	72148	14	AC090867 AC090867 Homo sapi
	317	16	55.2	73607	8	AL162740 AL162740 Human DNA
c	318	16	55.2	74001	8	AC058333 AC058333 Homo sapi
c	319	16	55.2	75295	6	AR659534 AR659534 Sequence
c	320	16	55.2	75296	6	AR659758 AR659758 Sequence
	321	16	55.2	75692	14	AC164790 AC164790 Bos tauru
c	322	16	55.2	75994	8	AC105417 AC105417 Homo sapi
	323	16	55.2	77134	8	AC146380 AC146380 Pan trogl
	324	16	55.2	77948	14	AC018982 AC018982 Homo sapi
c	325	16	55.2	78021	14	AC023096 AC023096 Homo sapi
	326	16	55.2	78167	8	AC092028 AC092028 Homo sapi
c	327	16	55.2	78167	8	AC092028 AC092028 Homo sapi
	328	16	55.2	78200	8	AP001866 AP001866 Homo sapi
c	329	16	55.2	78746	8	AC024289 AC024289 Homo sapi
c	330	16	55.2	79057	14	AC011735 AC011735 Homo sapi
	331	16	55.2	79419	8	HSJ324O17 AL110115 Human DNA
	332	16	55.2	80050	8	AC124833 AC124833 Homo sapi
c	333	16	55.2	80062	14	AC022560 AC022560 Homo sapi
c	334	16	55.2	80133	14	AC021347 AC021347 Homo sapi
c	335	16	55.2	80142	8	AL645465 AL645465 Human DNA
c	336	16	55.2	80343	8	HS474I12 AL022333 Human DNA
c	337	16	55.2	80364	8	CR388202 CR388202 Human DNA
	338	16	55.2	81231	8	AC003044 AC003044 Homo sapi
	339	16	55.2	81410	8	AC073594 AC073594 Homo sapi
c	340	16	55.2	81874	8	HS931E15 AL023575 Human DNA
	341	16	55.2	82011	8	AF209502 AF209502 Homo sapi
	342	16	55.2	82406	8	AC113430 AC113430 Homo sapi
	343	16	55.2	83352	8	AL139387 AL139387 Human DNA
	344	16	55.2	83661	8	AP001439 AP001439 Homo sapi
c	345	16	55.2	83675	8	AC004830 AC004830 Homo sapi
c	346	16	55.2	83840	8	HSJ148H17 AL109825 Human DNA
c	347	16	55.2	84114	8	AC003006 AC003006 Human DNA
c	348	16	55.2	84115	8	AC013468 AC013468 Homo sapi
c	349	16	55.2	84307	14	AC025206 AC025206 Homo sapi
	350	16	55.2	84661	8	AL353695 AL353695 Human DNA
	351	16	55.2	85144	8	AC125621 AC125621 Homo sapi
	352	16	55.2	85148	8	AC022337 AC022337 Homo sapi
	353	16	55.2	85351	14	AL390072_4 Continuation (5 of
c	354	16	55.2	85368	8	AC073516 AC073516 Homo sapi
c	355	16	55.2	85491	8	CR759915 CR759915 Human DNA
c	356	16	55.2	85854	14	AC140655 AC140655 Homo sapi
c	357	16	55.2	86684	8	AC010605 AC010605 Homo sapi
	358	16	55.2	86999	8	AC124947 AC124947 Homo sapi
	359	16	55.2	87034	8	AC091871 AC091871 Homo sapi
	360	16	55.2	88359	8	AL590305 AL590305 Human DNA
	361	16	55.2	88665	14	AL162261 AL162261 Homo sapi
c	362	16	55.2	89507	8	AL357132 AL357132 Human DNA
	363	16	55.2	89922	8	AC135584 AC135584 Homo sapi
	364	16	55.2	90079	14	AL731812 AL731812 Homo sapi
	365	16	55.2	90175	8	HSDJ858B6 AL118511 Human DNA
	366	16	55.2	91503	8	AL451061 AL451061 Human DNA
c	367	16	55.2	91822	8	AL160268 AL160268 Human DNA
	368	16	55.2	91830	8	AL365436 AL365436 Human DNA
	369	16	55.2	91842	8	AC008392 AC008392 Homo sapi
c	370	16	55.2	92059	8	BS000218 BS000218 Pan trogl
	371	16	55.2	92440	8	BX255934 BX255934 Human DNA
	372	16	55.2	92514	8	AL139148 AL139148 Human DNA
	373	16	55.2	92975	8	AC002369 AC002369 Homo sapi
c	374	16	55.2	93511	8	BX927132 BX927132 Human DNA
	375	16	55.2	93826	8	BX284668 BX284668 Human DNA
	376	16	55.2	94422	8	HSJ575L21 AL096841 Human DNA
	377	16	55.2	94793	8	AC006320 AC006320 Homo sapi
c	378	16	55.2	94793	8	AC006320 AC006320 Homo sapi
c	379	16	55.2	94826	14	AC146828 AC146828 Carollia
	380	16	55.2	96106	8	AC074119 AC074119 Homo sapi
	381	16	55.2	96146	8	AC095052 AC095052 Homo sapi
c	382	16	55.2	96153	8	AC092651 AC092651 Homo sapi
	383	16	55.2	96371	8	AF179633 AF179633 Homo sapi

c 384	16	55.2	97603	5	BX511193	BX511193 Zebrafish
c 385	16	55.2	97608	8	AL353586	AL353586 Human DNA
386	16	55.2	97693	8	AC002383	AC002383 Homo sapi
c 387	16	55.2	97757	8	AC092402	AC092402 Homo sapi
c 388	16	55.2	97835	8	HS292E10	Z93930 Human DNA s
389	16	55.2	98390	8	AL356986	AL356986 Human DNA
c 390	16	55.2	99322	14	AC046184	AC046184 Homo sapi
391	16	55.2	99452	8	AC096775	AC096775 Homo sapi
392	16	55.2	99886	8	AL135841	AL135841 Human DNA
c 393	16	55.2	99966	8	AP001606	AP001606 Homo sapi
394	16	55.2	100000	8	AP000143	AP000143 Homo sapi
395	16	55.2	100110	8	AC011506	AC011506 Homo sapi
c 396	16	55.2	100262	8	AL592211	AL592211 Human DNA
c 397	16	55.2	100496	8	HS970D1	AL021069 Human DNA
398	16	55.2	100541	8	DQ090940	DQ090940 Homo sapi
399	16	55.2	100634	8	AP001594	AP001594 Homo sapi
c 400	16	55.2	100711	8	AY623109	AY623109 Homo sapi
c 401	16	55.2	101083	14	AC139010	AC139010 Homo sapi
c 402	16	55.2	101629	14	AC021502	AC021502 Homo sapi
403	16	55.2	102086	8	AL512447	AL512447 Human DNA
c 404	16	55.2	102092	14	AC138139	AC138139 Homo sapi
c 405	16	55.2	102137	8	AL391868	AL391868 Human DNA
c 406	16	55.2	102196	8	CR388415	CR388415 Human DNA
c 407	16	55.2	102313	8	AL603749	AL603749 Human DNA
c 408	16	55.2	102326	8	AC087407	AC087407 Homo sapi
409	16	55.2	102509	8	AL953897	AL953897 Human DNA
c 410	16	55.2	102539	8	AC135895	AC135895 Homo sapi
c 411	16	55.2	102812	14	AL158215	AL158215 Homo sapi
412	16	55.2	103023	8	AP000747	AP000747 Homo sapi
c 413	16	55.2	103177	8	AP000841	AP000841 Homo sapi
414	16	55.2	103822	8	AL353639	AL353639 Human DNA
c 415	16	55.2	104002	8	HS388N15	Z99571 Human DNA s
416	16	55.2	104148	8	AC006146	AC006146 Homo sapi
417	16	55.2	104552	8	AL137159	AL137159 Human DNA
c 418	16	55.2	104600	8	AP001166	AP001166 Homo sapi
c 419	16	55.2	104660	8	AP000347	AP000347 Homo sapi
420	16	55.2	104776	8	AC022424	AC022424 Homo sapi
421	16	55.2	104810	8	AC003012	AC003012 Human PAC
422	16	55.2	104818	14	AC026690	AC026690 Homo sapi
423	16	55.2	106063	14	AL133290	AL133290 Homo sapi
c 424	16	55.2	106216	8	AC061997	AC061997 Homo sapi
c 425	16	55.2	106239	8	AL607122	AL607122 Human DNA
c 426	16	55.2	106526	8	AC104772	AC104772 Homo sapi
c 427	16	55.2	106735	14	AC011381	AC011381 Homo sapi
428	16	55.2	107260	8	AL139352	AL139352 Human DNA
429	16	55.2	107336	8	AC011648	AC011648 Homo sapi
c 430	16	55.2	107714	14	AL356306	AL356306 Homo sapi
431	16	55.2	108086	8	HS220J16	AL023798 Human DNA
c 432	16	55.2	108315	8	HS41P2	AL049748 Human DNA
433	16	55.2	109081	8	AL358074	AL358074 Human DNA
c 434	16	55.2	109445	8	AC008946	AC008946 Homo sapi
c 435	16	55.2	110000	1	CP000089_35	Continuation (36 o
436	16	55.2	110000	8	BA000025_01	Continuation (2 of
437	16	55.2	110000	14	AC003656_1	Continuation (2 of
c 438	16	55.2	110000	14	AC003656_1	Continuation (2 of
c 439	16	55.2	110000	14	AC009727_1	Continuation (2 of
440	16	55.2	110000	14	AC098270_1	Continuation (2 of
441	16	55.2	110000	14	AC112133_1	Continuation (2 of
c 442	16	55.2	110000	14	AC118411_2	Continuation (3 of
c 443	16	55.2	110000	14	AC121437_1	Continuation (2 of
444	16	55.2	110000	14	AC156265_13	Continuation (14 o
445	16	55.2	110000	14	AC156265_14	Continuation (15 o
c 446	16	55.2	110000	14	AL353694_1	Continuation (2 of
c 447	16	55.2	110000	14	AL353694_2	Continuation (3 of
448	16	55.2	110000	14	AL390072_3	Continuation (4 of
c 449	16	55.2	110000	14	AL713858_0	AL713858 Homo sapi
450	16	55.2	110000	14	BX324168_0	BX324168 Homo sapi
451	16	55.2	110000	14	CR753903_1	Continuation (2 of
452	16	55.2	110000	14	CR759862_0	CR759862 Homo sapi
c 453	16	55.2	110000	14	CR759862_2	Continuation (3 of
c 454	16	55.2	110000	14	CR936358_4	Continuation (5 of
455	16	55.2	110000	15	AP008208_207	Continuation (208
c 456	16	55.2	110077	8	AC011346	AC011346 Homo sapi
c 457	16	55.2	110401	14	AC104649	AC104649 Homo sapi
458	16	55.2	110554	8	AC107376	AC107376 Homo sapi
459	16	55.2	110554	8	AF165146	AF165146 Homo sapi
c 460	16	55.2	111084	8	AC006486	AC006486 Homo sapi
c 461	16	55.2	111178	8	AL136129	AL136129 Human DNA
c 462	16	55.2	111566	8	AL357314	AL357314 Human DNA
c 463	16	55.2	111633	8	BX248093	BX248093 Human DNA
c 464	16	55.2	111873	8	AC073275	AC073275 Homo sapi
c 465	16	55.2	112058	14	AC055112	AC055112 Homo sapi
c 466	16	55.2	112106	8	AC105142	AC105142 Homo sapi
467	16	55.2	112224	8	HS423B22	AL034379 Human DNA
c 468	16	55.2	112672	8	AC010455	AC010455 Homo sapi
469	16	55.2	113024	14	AL445260	AL445260 Homo sapi
c 470	16	55.2	113194	14	AF301236	AF301236 Homo sapi
c 471	16	55.2	113582	8	AL671762	AL671762 Human DNA
472	16	55.2	114233	8	AP000783	AP000783 Homo sapi
c 473	16	55.2	114694	8	AL355500	AL355500 Human DNA
474	16	55.2	114771	8	HS569D19	AL022334 Human DNA
c 475	16	55.2	115281	8	AC013815	AC013815 Homo sapi
c 476	16	55.2	115710	14	AL139015	AL139015 Homo sapi
c 477	16	55.2	117487	14	AP002335	AP002335 Homo sapi
c 478	16	55.2	117564	14	HSDJ160H5	AL096873 Homo sapi
c 479	16	55.2	118047	8	HSDJ513M9	AL049631 Human DNA
480	16	55.2	118269	8	HSA541I12	AL121868 Human DNA
c 481	16	55.2	118741	8	AC005319	AC005319 Human Chr
482	16	55.2	118763	8	AC114484	AC114484 Homo sapi
c 483	16	55.2	118873	8	HS1184F4	AL034550 Human DNA
484	16	55.2	119122	8	AC129906	AC129906 Homo sapi
485	16	55.2	119147	8	AC098873	AC098873 Homo sapi
c 486	16	55.2	119347	8	AC020934	AC020934 Homo sapi
c 487	16	55.2	119555	8	BX004827	BX004827 Human DNA
488	16	55.2	119884	8	AL139814	AL139814 Human DNA
c 489	16	55.2	120088	8	AC004865	AC004865 Homo sapi
c 490	16	55.2	120510	8	AL356457	AL356457 Human DNA
491	16	55.2	121377	15	AP004038	AP004038 Oryza sat
492	16	55.2	121419	5	BX324226	BX324226 Zebrafish
493	16	55.2	121597	8	AP000090	AP000090 Homo sapi
c 494	16	55.2	122176	14	AL358233	AL358233 Homo sapi
c 495	16	55.2	122934	8	AC092898	AC092898 Homo sapi
496	16	55.2	123253	8	AC004971	AC004971 Homo sapi
497	16	55.2	123631	8	HS22F01	AL109967 Homo sapi
498	16	55.2	123830	8	HS500L14	AL023583 Human DNA
c 499	16	55.2	124152	9	AC147048	AC147048 Mus muscu
500	16	55.2	124536	14	AC149107	AC149107 Papio anu

ALIGNMENTS

```
RESULT 1
AX300008
LOCUS       AX300008               2145 bp    DNA     linear   PAT 26-NOV-2001
DEFINITION  Sequence 13 from Patent WO0166740.
ACCESSION   AX300008
VERSION     AX300008.1  GI:17129479
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Eaton,D.L., Fong,S., Goddard,A., Godowski,P.J., Grimaldi,C.J.,
            Gurney,A.L., Tumas,D., Watanabe,C.K., Wood,W.I. and Zhang,Z.
  TITLE     Compositions and methods for the treatment of immune related
            diseases
  JOURNAL   Patent: WO 0166740-A 13 13-SEP-2001;
            Genentech, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2145
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
```

```
ORIGIN

  Query Match          65.5%;  Score 19;  DB 6;  Length 2145;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          11 GCTCAGGCATGAGCCAGCA 29
               |||||||||||||||||||
Db           1 GCTCAGGCATGAGCCAGCA 19


RESULT 2
AC069487/c
LOCUS       AC069487               170301 bp    DNA     linear   HTG 30-JUN-2000
DEFINITION  Homo sapiens chromosome 17 clone RP11-651P17 map 17, WORKING DRAFT
            SEQUENCE, 14 unordered pieces.
ACCESSION   AC069487
VERSION     AC069487.2  GI:8844144
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 170301)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone RP11-651P17
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 170301)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jun 30, 2000 this sequence version replaced gi:8137301.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L10468
                Center clone name: 651_P_17
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 162206 bases at least Q40
                Consensus quality: 166475 bases at least Q30
                Consensus quality: 168143 bases at least Q20
                Insert size: 169001; sum-of-contigs
                Quality coverage: 4.8 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 14 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     1111: contig of 1111 bp in length
            *    1112     1211: gap of 100 bp
            *    1212     2636: contig of 1425 bp in length
            *    2637     2736: gap of 100 bp
            *    2737     4529: contig of 1793 bp in length
            *    4530     4629: gap of 100 bp
            *    4630     6040: contig of 1411 bp in length
            *    6041     6140: gap of 100 bp
            *    6141     7267: contig of 1127 bp in length
            *    7268     7367: gap of 100 bp
            *    7368    10414: contig of 3047 bp in length
            *   10415    10514: gap of 100 bp
            *   10515    16074: contig of 5560 bp in length
            *   16075    16174: gap of 100 bp
            *   16175    24878: contig of 8704 bp in length
            *   24879    24978: gap of 100 bp
            *   24979    35974: contig of 10996 bp in length
            *   35975    36074: gap of 100 bp
            *   36075    51535: contig of 15461 bp in length
            *   51536    51635: gap of 100 bp
            *   51636    70368: contig of 18733 bp in length
            *   70369    70468: gap of 100 bp
            *   70469    90946: contig of 20478 bp in length
            *   90947    91046: gap of 100 bp
            *   91047   129158: contig of 38112 bp in length
            *  129159   129258: gap of 100 bp
            *  129259   170301: contig of 41043 bp in length.
FEATURES             Location/Qualifiers
     source          1. .170301
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP11-651P17"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1111
                     /note="assembly_fragment"
     gap             1112. .1211
                     /estimated_length=100
     misc_feature    1212. .2636
                     /note="assembly_fragment"
     gap             2637. .2736
                     /estimated_length=100
     misc_feature    2737. .4529
                     /note="assembly_fragment"
     gap             4530. .4629
                     /estimated_length=100
     misc_feature    4630. .6040
                     /note="assembly_fragment"
     gap             6041. .6140
                     /estimated_length=100
     misc_feature    6141. .7267
                     /note="assembly_fragment"
     gap             7268. .7367
                     /estimated_length=100
     misc_feature    7368. .10414
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     gap             10415. .10514
                     /estimated_length=100
     misc_feature    10515. .16074
```

```
                    /note="assembly_fragment"
     gap            16075. .16174
                    /estimated_length=100
     misc_feature   16175. .24878
                    /note="assembly_fragment
                    clone_end:T7
                    vector_side:left"
     gap            24879. .24978
                    /estimated_length=100
     misc_feature   24979. .35974
                    /note="assembly_fragment"
     gap            35975. .36074
                    /estimated_length=100
     misc_feature   36075. .51535
                    /note="assembly_fragment"
     gap            51536. .51635
                    /estimated_length=100
     misc_feature   51636. .70368
                    /note="assembly_fragment"
     gap            70369. .70468
                    /estimated_length=100
     misc_feature   70469. .90946
                    /note="assembly_fragment"
     gap            90947. .91046
                    /estimated_length=100
     misc_feature   91047. .129158
                    /note="assembly_fragment"
     gap            129159. .129258
                    /estimated_length=100
     misc_feature   129259. .170301
                    /note="assembly_fragment"
ORIGIN

  Query Match             65.5%;  Score 19;  DB 14;  Length 170301;
  Best Local Similarity   100.0%;  Pred. No. 0.58;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 GCTCAGGCATGAGCCAGCA 29
              |||||||||||||||||||
Db     119787 GCTCAGGCATGAGCCAGCA 119769


RESULT 3
HS170A21
LOCUS       HS170A21              176947 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP1-170A21 on chromosome 22, complete
            sequence.
ACCESSION   Z82189
VERSION     Z82189.1  GI:3164069
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 176947)
  AUTHORS   Connor,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On May 28, 1998 this sequence version replaced gi:1772936.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 22, constructed by the Sanger Centre Chromosome 22
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr22
            RP1-170A21 is from the library RPCI-1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            ------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            -------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .176947
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIP704A21170"
                     /db_xref="taxon:9606"
                     /chromosome="22"
                     /clone="RP1-170A21"
                     /clone_lib="RPCI-1"
     gene            complement(join(140727. .140832,140912. .141028))
                     /locus_tag="RP1-170A21.4-004"
     mRNA            complement(join(140727. .140832,140912. .141028))
                     /locus_tag="RP1-170A21.4-004"
ORIGIN

  Query Match             65.5%;  Score 19;  DB 8;  Length 176947;
  Best Local Similarity   100.0%;  Pred. No. 0.58;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 GCTCAGGCATGAGCCAGCA 29
              |||||||||||||||||||
Db      56975 GCTCAGGCATGAGCCAGCA 56993


RESULT 4
AC078932/c
LOCUS       AC078932              238521 bp    DNA     linear   HTG 11-AUG-2000
DEFINITION  Mus musculus strain C57BL6/J chromosome 6 clone RP23-11P24, WORKING
            DRAFT SEQUENCE, 21 unordered pieces.
ACCESSION   AC078932
VERSION     AC078932.1  GI:9795568
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 238521)
  AUTHORS   Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W.,
            Bouffard,G.G., Dietrich,N.L., Eagle,W.O., Gupta,J., Ho,S.-L.,
            Huang,M.C., Idol,J., Lee-Lin,S.-Q., Maduro,Q.L., Maduro,V.B.,
            Mastrian,S.D., McCloskey,J.C., McDowell,J., Ojodu,M.A., Pearson,R.,
            Stantripop,S., Summers,T.J., Thomas,J.W., Thomas,P.J.,
            Tiongson,E.E., Touchman,J.W., Tran,J.T., Vogt,J.L., Walker,M.A.,
            Wetherby,K.D. and Green,E.D.
  TITLE     NISC Mouse Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 238521)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-AUG-2000) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT     -------------- Genome Center
                 Center: NIH Intramural Sequencing Center
                 Center code: NISC
                 Web site: http://www.nisc.nih.gov
                 Contact: nisc_mouse@nhgri.nih.gov
            -------------- Project Information
```

```
            Center project name: tf
            Center clone name: 011P24
          ------------- Summary Statistics
            Sequencing vector: plasmid; n/a; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 225819 bases at least Q40
            Consensus quality: 229481 bases at least Q30
            Consensus quality: 231007 bases at least Q20
            Insert size: 216000; agarose-fp
            Insert size: 248000; pulse-field-gel
            Insert size: 236521; sum-of-contigs
            Quality coverage: 5.18x in Q20 bases; agarose-fp
            Quality coverage: 4.51x in Q20 bases; pulse-field-gel
            Quality coverage: 4.73x in Q20 bases; sum-of-contigs
          --------------.
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 21 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *        1     3197: contig of 3197 bp in length
          *     3198     3297: gap of unknown length
          *     3298     5529: contig of 2232 bp in length
          *     5530     5629: gap of unknown length
          *     5630     9534: contig of 3905 bp in length
          *     9535     9634: gap of unknown length
          *     9635    12112: contig of 2478 bp in length
          *    12113    12212: gap of unknown length
          *    12213    17291: contig of 5079 bp in length
          *    17292    17391: gap of unknown length
          *    17392    23721: contig of 6330 bp in length
          *    23722    23821: gap of unknown length
          *    23822    29985: contig of 6164 bp in length
          *    29986    30085: gap of unknown length
          *    30086    34329: contig of 4244 bp in length
          *    34330    34429: gap of unknown length
          *    34430    40695: contig of 6266 bp in length
          *    40696    40795: gap of unknown length
          *    40796    45075: contig of 4280 bp in length
          *    45076    45175: gap of unknown length
          *    45176    53129: contig of 7954 bp in length
          *    53130    53229: gap of unknown length
          *    53230    57922: contig of 4693 bp in length
          *    57923    58022: gap of unknown length
          *    58023    67177: contig of 9155 bp in length
          *    67178    67277: gap of unknown length
          *    67278    77163: contig of 9886 bp in length
          *    77164    77263: gap of unknown length
          *    77264    86276: contig of 9013 bp in length
          *    86277    86376: gap of unknown length
          *    86377   101681: contig of 15305 bp in length
          *   101682   101781: gap of unknown length
          *   101782   118067: contig of 16286 bp in length
          *   118068   118167: gap of unknown length
          *   118168   134707: contig of 16540 bp in length
          *   134708   134807: gap of unknown length
          *   134808   160231: contig of 25424 bp in length
          *   160232   160331: gap of unknown length
          *   160332   194122: contig of 33791 bp in length
          *   194123   194222: gap of unknown length
          *   194223   238521: contig of 44299 bp in length.
FEATURES             Location/Qualifiers
     source          1. .238521
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL6/J"
                     /db_xref="taxon:10090"
                     /chromosome="6"
                     /clone="RP23-11P24"
                     /clone_lib="RPCI mouse BAC library 23"
     misc_feature    1. .3197
                     /note="assembly_fragment"
     gap             3198. .3297
                     /estimated_length=unknown
     misc_feature    3298. .5529
                     /note="assembly_fragment"
     gap             5530. .5629
                     /estimated_length=unknown
     misc_feature    5630. .9534
                     /note="assembly_fragment"
     gap             9535. .9634
                     /estimated_length=unknown
     misc_feature    9635. .12112
                     /note="assembly_fragment"
     gap             12113. .12212
                     /estimated_length=unknown
     misc_feature    12213. .17291
                     /note="assembly_fragment"
     gap             17292. .17391
                     /estimated_length=unknown
     misc_feature    17392. .23721
                     /note="assembly_fragment"
     gap             23722. .23821
                     /estimated_length=unknown
     misc_feature    23822. .29985
                     /note="assembly_fragment"
     gap             29986. .30085
                     /estimated_length=unknown
     misc_feature    30086. .34329
                     /note="assembly_fragment"
     gap             34330. .34429
                     /estimated_length=unknown
     misc_feature    34430. .40695
                     /note="assembly_fragment"
     gap             40696. .40795
                     /estimated_length=unknown
     misc_feature    40796. .45075
                     /note="assembly_fragment"
     gap             45076. .45175
                     /estimated_length=unknown
     misc_feature    45176. .53129
                     /note="assembly_fragment"
     gap             53130. .53229
                     /estimated_length=unknown
     misc_feature    53230. .57922
                     /note="assembly_fragment"
     gap             57923. .58022
                     /estimated_length=unknown
     misc_feature    58023. .67177
                     /note="assembly_fragment"
     gap             67178. .67277
                     /estimated_length=unknown
     misc_feature    67278. .77163
                     /note="assembly_fragment"
     gap             77164. .77263
                     /estimated_length=unknown
     misc_feature    77264. .86276
                     /note="assembly_fragment"
     gap             86277. .86376
                     /estimated_length=unknown
     misc_feature    86377. .101681
                     /note="assembly_fragment"
     gap             101682. .101781
                     /estimated_length=unknown
     misc_feature    101782. .118067
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     gap             118068. .118167
                     /estimated_length=unknown
     misc_feature    118168. .134707
                     /note="assembly_fragment"
```

```
     gap             134708. .134807
                     /estimated_length=unknown
     misc_feature    134808. .160231
                     /note="assembly_fragment"
     gap             160232. .160331
                     /estimated_length=unknown
     misc_feature    160332. .194122
                     /note="assembly_fragment"
     gap             194123. .194222
                     /estimated_length=unknown
     misc_feature    194223. .238521
                     /note="assembly_fragment"
ORIGIN

  Query Match             65.5%;  Score 19;  DB 14;  Length 238521;
  Best Local Similarity   100.0%;  Pred. No. 0.56;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAGCCA 26
              ||||||||||||||||||
Db     130261 GGAGCTCAGGCATGAGCCA 130243


RESULT 5
AC153817
LOCUS       AC153817               247508 bp    DNA     linear   ROD 30-APR-2005
DEFINITION  Mus musculus  BAC RP23-11P24 (Roswell Park Cancer Institute
            (C57BL/6J Female) Mouse BAC Library) complete sequence.
ACCESSION   AC153817
VERSION     AC153817.6  GI:62988373
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 247508)
  AUTHORS   Muzny,D., Adams,C., Agbai II,O., Allen,C., Alsbrooks,S., Archer,P.,
            Arredondo,H., Bandaranaike,D., Bangura,L., Beltran,B., Beltran,R.,
            Beraducci,A., Biswalo,K., Blyth,P., Bonham,H., Buhay,C., Burch,P.,
            Cadoree,I., Canada,A., Cardenas,V., Carter,K., Cavazos,I.,
            Chacko,J., Chahrour,M., Chavez,D., Chen,A., Chen,G., Chen,R.,
            Cheng,M.-T., Chu,J., Clerc,K., Cockrell,R., Coyle,M., Cree,A.,
            Curry,S., Dai,W., Davila,M.L., Davis,C., Davy-Carroll,L., De
            Anda,C., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H.,
            Donlin,J., McCauley,S., Dugan-Rocha,S., Dunn,A., Durbin,K.,
            Dziuda,D., Egan,A., Escotto,M., Espinosa,V., Eugene,C., Fa,M.,
            Fernandez,S., Fernando,P., Flagg,N., Forbes,L., Foster,P.,
            Fowler,G., Fu,Q., Fuh,E., Garcia,A., Garcia,R., Garner,T.,
            Gaskin,C., Gench,S., Ghose,S., Gill,R., Gonzalez,D.,
            Gonzalez-Garay,M., Guevara,W., Holder,M., Haaland,W., Haeberlen,K.,
            Hall,B., Hamid,H., Hamilton,K., Harbes,B., Harris,R., Havlak,P.,
            Hawes,A., Hawkins,E., Hayes,S., Hemphill,L., Hernandez,J.,
            Hines,S., Hitchens,M., Hodgson,A., Hogues,M., Hollins,B.,
            Howell,L.T., Hulyk,S., Hume,J., Imo,K., Jackson,A., Jackson,L.,
            Jacob,L., Jiang,H., Johnson,B., Johnson,R., Kalafus,K., Kelly,S.,
            Keys,T., Khan,Z., King,L., Kovar,C., Kowis,A., Kowis,C., Lara,F.,
            Leal,S., Lee,K., Lee,S., LeGall,F.I., Lemon,S., Lewis,L., Li,B.,
            Li,Y., Li,Z., Linnell,M., Liu,W., Liu,Y.-S., Liu,Y., Liyanage,D.,
            London,P., Lopez,J., Lorensuhewa,L., Lozado,R., Luk,T., Madu,R.,
            Maheshwari,M., Mahoney,C., Malloy,K., Mansouri,D., Martinez,E.,
            McClelland,H., McPherson,J., Mercadao,C., Metzker,M.,
            Milosavljevic,A., Minja,E., Morgan,M., Morris,S., Munidasa,M.,
            Murray,D., Nazarith,L., Ngo,D., Nguyen,N., Norwig-Eastaugh,E.,
            Nott,A., Nwaokelemeh,O., Obregon,M., Ochi-Okorie,C., Odeh,E.,
            Okwuonu,G., Okwuonu,K., Parker,D., Pasternak,S., Patel,B.,
            Patel,V., Paul,H., Perez,A., Perez,L., Petrosino,J., Pham,T.,
            Primus,E., Pu,L.-L., Puazo,M., Qin,X., Quinn,A., Quiroz,J.,
            Rabata,D., Rachlin,E., Reigh,R., Ren,Y., Reuter,M., Richards,S.,
            Rives,C., Rodriguez,F., Rojas,A., Ruiz,S.J., Sana,M., Sanders,W.,
            Santibanez,J., Santos,R., Savery,G., Scherer,S., Shen,H., Shen,Y.,
            Sisson,I., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R.,
            Svatek,A., Taylor,E., Taylor,T., Thomas,N., Thorn,R., Thornton,R.,
            Trejos,Z., Usmani,K., Vargo,C., Verduzco,D., Villasana,D., Virk,D.,
            Volkov,A., Waldron,L., Walker,B., Wang,Q., Wang,S., Warren,J.,
            Wei,X., Wheeler,D., Williams,G., Williams,R., Worley,K., Wright,R.,
            Wu,J., Yakub,S., Yan,K., Yuan,Y., Yu,F., Zhang,J., Zhang,L.,
            Zhang,Z., Zhou,J., Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 247508)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-DEC-2004) Human Genome Sequencing Center, Baylor
            College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 247508)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2005) Human Genome Sequencing Center, Baylor
            College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 247508)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-APR-2005) Human Genome Sequencing Center, Baylor
            College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Apr 30, 2005 this sequence version replaced gi:61889150.
            Sequencing is completed to a minimum standard of double strand
            coverage with a minimum of 2 clones and 2 reads with no ambiguities
            or 2 chemistries with a minimum of 2 clones and 3 reads with no
            ambiguities. If the sequence quality does not meet this standard,
            it will be indicated in the annotation.

            The repeat regions shown were identified using RepeatMasker by
            Adrian Smit.

            Sequence similarities were identified using Powerblast by Jinghui
            Zhang.

            Exon/Intron boundaries of identified genes were chosen if there
            were canonical splice junctions that maintained sequence continuity
            across the splice junctions.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Drafting Center Code: BCM
                Contact: hgsc-help@bcm.tmc.edu.
FEATURES             Location/Qualifiers
     source          1. .247508
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="RP23-11P24"
     misc_feature    complement(1. .51751)
                     /note="overlaps bases 1. .51751 of clone AC155915"
                     /function="clone overlap"
     misc_feature    38649. .38693
                     /note="Sized by Restriction Digest and PCR
                      244 bp"
                     /function="unresolved tandem repeat"
     misc_feature    172994. .173089
                     /note="Sized by Restriction Digest and PCR
                      145 bp , force join @ 173007 bp - 173026 bp"
                     /function="unresolved tandem repeat"
ORIGIN

  Query Match             65.5%;  Score 19;  DB 9;  Length 247508;
  Best Local Similarity   100.0%;  Pred. No. 0.55;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAGCCA 26
              ||||||||||||||||||
Db      99449 GGAGCTCAGGCATGAGCCA 99467


RESULT 6
```

```
AF181894/c
LOCUS       AF181894               2300 bp    mRNA    linear   VRT 20-JUN-2000
DEFINITION  Taricha granulosa CB1 cannabinoid receptor mRNA, complete cds.
ACCESSION   AF181894
VERSION     AF181894.1  GI:5917766
KEYWORDS    .
SOURCE      Taricha granulosa (rough-skinned newt)
  ORGANISM  Taricha granulosa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;
            Taricha.
REFERENCE   1  (bases 1 to 2300)
  AUTHORS   Soderstrom,K., Leid,M., Moore,F.L. and Murray,T.F.
  TITLE     Behaviroal, pharmacological, and molecular characterization of an
            amphibian cannabinoid receptor
  JOURNAL   J. Neurochem. 75 (1), 413-423 (2000)
   PUBMED   10854287
REFERENCE   2  (bases 1 to 2300)
  AUTHORS   Soderstrom,K., Moore,F.L., Leid,M. and Murray,T.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-AUG-1999) Psychology, Florida State University,
            KRB329, Tallahassee, FL 32306-1270, USA
FEATURES             Location/Qualifiers
     source          1. .2300
                     /organism="Taricha granulosa"
                     /mol_type="mRNA"
                     /db_xref="taxon:8321"
     CDS             645. .2066
                     /note="G-protein-coupled cannabinoid receptor;
                     CNS-associated"
                     /codon_start=1
                     /product="CB1 cannabinoid receptor"
                     /protein_id="AAD56029.1"
                     /db_xref="GI:5917767"
                     /translation="MKSILDGLADTTFRTITTDLLYMGSNDVQYEDTKGEMASKLGYF
                     PQKLPLSSFRRDHSPDKMTIGDDNLLSFYPLDQFNVTEPFNRSVSTFKENDDNLKCGE
                     NFMDMECFMILTASQQLIIAVLSLTLGTFTVLENFLVLCVILQSRTLRCRPSYHFIGS
                     LAVADLLGSVIFVYSFLDFHVFHRKDSSNVFLFKLGGVTASFTASVGSLFLTAIDRYI
                     SIHRPLAYKRIVTRTKAVIAFCVMWTIAIIIAVLPLLGWNCKKLKSVCSDIFPLIDEN
                     YLMFWIGVTSILLLFIVYAYVYILWKAHSHAVRMLQRGTQKSIIIHTSEDGKVQITRP
                     EQTRMDIRLAKTLVLILVVLIICWGPLLAIMVYDVFGKMNNPIKTVFAFCSMLCLMDS
                     TVNPIIYALRSQDLRHAFLEQCPPCEGTSQPLDNSMESDCQHRHGNNAGNVHRAAENC
                     IKSTVKIAKVTMSVSTETSGEAV"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 5;  Length 2300;
  Best Local Similarity   100.0%;  Pred. No. 4.6;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CTCTCTGGAGCTCAGGCA 19
              ||||||||||||||||||
Db        482 CTCTCTGGAGCTCAGGCA 465


RESULT 7
BX539329
LOCUS       BX539329               34713 bp    DNA     linear   HTG 06-JUL-2003
DEFINITION  Homo sapiens chromosome 20 clone XX-R19IB10_1, WORKING DRAFT
            SEQUENCE, 6 unordered pieces.
ACCESSION   BX539329
VERSION     BX539329.1  GI:32134761
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 34713)
  AUTHORS   Burton,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUL-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     ---------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            ---------- Project Information
            Center project name: fyR19IB10_1
            ---------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 32878 bases at least Q40
            Consensus quality: 33404 bases at least Q30
            Consensus quality: 33685 bases at least Q20
            Insert size: 34213; sum-of-contigs
            Insert size: 42757; 5.5% error; agarose-fp
            Quality coverage: 4.85x in Q20 bases; sum-of-contigs Quality
            coverage: 4.33x in Q20 bases; agarose-fp
            ----------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2006: contig of 2006 bp in length
            *     2007     2106: gap of 100 bp
            *     2107     9437: contig of 7331 bp in length
            *     9438     9537: gap of 100 bp
            *     9538    13772: contig of 4235 bp in length
            *    13773    13872: gap of 100 bp
            *    13873    20486: contig of 6614 bp in length
            *    20487    20586: gap of 100 bp
            *    20587    23320: contig of 2734 bp in length
            *    23321    23420: gap of 100 bp
            *    23421    34713: contig of 11293 bp in length.
FEATURES             Location/Qualifiers
     source          1. .34713
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /clone="XX-R19IB10_1"
                     /clone_lib="Human fosmid library"
     misc_feature    1. .2006
                     /note="assembly_fragment:00386
                     fragment_chain:1
                     clone_end:SP6
                     vector_side:left"
     misc_feature    2107. .9437
                     /note="assembly_fragment:00269
                     fragment_chain:1"
     misc_feature    9538. .13772
                     /note="assembly_fragment:00159
                     fragment_chain:1"
     misc_feature    13873. .20486
                     /note="assembly_fragment:00353
                     fragment_chain:1"
     misc_feature    20587. .23320
                     /note="assembly_fragment:00292
                     fragment_chain:1"
     misc_feature    23421. .34713
                     /note="assembly_fragment:00247"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 14;  Length 34713;
  Best Local Similarity   100.0%;  Pred. No. 3.1;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         12 CTCAGGCATGAGCCAGCA 29
              ||||||||||||||||||
Db      10963 CTCAGGCATGAGCCAGCA 10980
```

```
RESULT 8
AL450465
LOCUS       AL450465               47139 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP3-493M24 on chromosome 20 Contains
            ESTs and GSSs, complete sequence.
ACCESSION   AL450465
VERSION     AL450465.12  GI:12584473
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 47139)
  AUTHORS   Tracey,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jan 28, 2001 this sequence version replaced gi:12581069.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 20, constructed by the Sanger Centre Chromosome 20
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr20
            RP3-493M24 is from the library RPCI-3 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .47139
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIP704M24493"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /clone="RP3-493M24"
                     /clone_lib="RPCI-3"
     misc_feature    1
                     /note="Clone_left_end: RP3-493M24"
     misc_feature    47040
                     /note="Clone_left_end: RP5-996C2"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 8;  Length 47139;
  Best Local Similarity   100.0%;  Pred. No. 2.9;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          12 CTCAGGCATGAGCCAGCA 29
               ||||||||||||||||||
Db       10067 CTCAGGCATGAGCCAGCA 10084


RESULT 9
AL928544
LOCUS       AL928544               126329 bp    DNA     linear   ROD 07-JUL-2005
DEFINITION  Mouse DNA sequence from clone RP23-358C5 on chromosome 2, complete
            sequence.
ACCESSION   AL928544
VERSION     AL928544.5  GI:70562999
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 126329)
  AUTHORS   North,P., Leaves,N., Greystrong,J., Coppola,M., Manjunath,S.,
            Russell,E., Smith,M., Strachan,G., Tofts,C., Boal,E., Cobley,V.,
            Hunter,G., Kimberley,C., Thomas,D., Cave-Berry,L., Weston,P. and
            Botcherby,M.R.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUL-2005) Mouse Sequencing Group, HGMP-RC, Hinxton,
            Cambridge, CB10 1SB, UK. E-mail enquiries:- mrbotche@hgmp.mrc.ac.uk
            or pnorth@hgmp.mrc.ac.uk
  REMARK    HGMP-RC part of the UK Mouse Sequencing Consortium
COMMENT     On Jul 7, 2005 this sequence version replaced gi:40789205.
            -------------- Genome Center
            Center: UK Medical Research Council
            Center code: UK-MRC
            Web site: http://mrcseq.har.mrc.ac.uk
            Contact: mouseq@har.mrc.ac.uk
            --------------
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-358C5 is
            from the RPCI-23 Mouse BAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            Sequence from the Mouse Genome Sequencing Consortium whole genome
            shotgun may have been used to confirm this sequence.  Sequence data
            from the whole genome shotgun alone has only been used where it has
            a phred quality of at least 30.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .126329
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="2"
                     /clone="RP23-358C5"
                     /clone_lib="RPCI-23"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 9;  Length 126329;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CTCTCTGGAGCTCAGGCA 19
              ||||||||||||||||||
Db      53991 CTCTCTGGAGCTCAGGCA 54008


RESULT 10
BX649367/c
LOCUS       BX649367               161289 bp    DNA     linear   VRT 07-NOV-2003
DEFINITION  Zebrafish DNA sequence from clone CH211-133F22 in linkage group 20,
            complete sequence.
```

```
ACCESSION   BX649367
VERSION     BX649367.6  GI:37650771
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 161289)
  AUTHORS   Babbage,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-NOV-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 10, 2003 this sequence version replaced gi:37518514.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml
            CH211-133F22 is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1.
FEATURES             Location/Qualifiers
     source          1. .161289
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-133F22"
                     /clone_lib="CHORI-211"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 5;  Length 161289;
  Best Local Similarity   100.0%;  Pred. No. 2.4;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 CTCTGGAGCTCAGGCATG 21
              ||||||||||||||||||
Db     145287 CTCTGGAGCTCAGGCATG 145270


RESULT 11
AL157935/c
LOCUS       AL157935              164201 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-203J24 on chromosome 9 Contains
            the 5' end of the ENG gene for endoglin (Osler-Rendu-Weber syndrome
            1) (END, ORW, HHT1, ORW1, CD105), the AK1 gene for adenylate kinase
            1, the ST6GALNAC6 gene for CMP-NeuAC:(beta)-N-acetylgalactosaminide
            (alpha)2,6-sialyltransferase member VI (ST6GALNACVI), the SIAT7D
            gene for sialyltransferase 7D
            ((alpha-N-acetylneuraminyl-2,3-beta-galactosyl-1,3)-N-acetyl
            galactosaminide alpha-2,6-sialyltransferase) (SIAT3C, ST6GALNAC4,
            ST6GALNACIV), the DPM2 gene for dolichyl-phosphate
            mannosyltransferase polypeptide 2, regulatory subunit (MGC21559),
            the gene for a novel protien containing FLJ00179, a novel gene and
            ten CpG islands, complete sequence.
ACCESSION   AL157935
VERSION     AL157935.28  GI:17221172
KEYWORDS    HTG; CD105; CpG island; DPM2; END; ENG; FLJ00179; HHT1; MGC21559;
            ORW; ORW1; SIAT3C; SIAT7D; ST6GALNAC4; ST6GALNAC6; ST6GALNACIV;
            ST6GALNACVI.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 164201)
  AUTHORS   Corby,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 30, 2001 this sequence version replaced gi:16944853.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr9
            RP11-203J24 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .164201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="9"
                     /clone="RP11-203J24"
                     /clone_lib="RPCI-11.1"
     misc_feature    1
                     /note="Clone_left_end: RP11-203J24"
     gene            join(complement(27622. .28089),complement(16427. .16578),
                     complement(3020. .3160),
                     complement(AL162586.26:78485. .78647),
                     complement(AL162586.26:77670. .77835),
                     complement(AL162586.26:77206. .77332),
                     complement(AL162586.26:76775. .76949),
                     complement(AL162586.26:76279. .76421),
                     complement(AL162586.26:71875. .72012),
                     complement(AL162586.26:71597. .71635),
```

```
                complement(AL162586.26:70691. .70807),
                complement(AL162586.26:70095. .70352),
                complement(AL162586.26:69124. .69178),
                complement(AL162586.26:67918. .68028),
                complement(AL162586.26:66987. .67781))
                /gene="ENG"
                /locus_tag="RP11-228B15.2-002"
     mRNA       join(complement(27622. .28089),complement(16427. .16578),
                complement(3020. .3160),
                complement(AL162586.26:78485. .78647),
                complement(AL162586.26:77670. .77835),
                complement(AL162586.26:77206. .77332),
                complement(AL162586.26:76775. .76949),
                complement(AL162586.26:76279. .76421),
                complement(AL162586.26:71875. .72012),
                complement(AL162586.26:71597. .71635),
                complement(AL162586.26:70691. .70807),
                complement(AL162586.26:70095. .70352),
                complement(AL162586.26:69124. .69178),
                complement(AL162586.26:67918. .68028),
                complement(AL162586.26:66987. .67781))
                /gene="ENG"
                /locus_tag="RP11-228B15.2-002"
                /product="endoglin (Osler-Rendu-Weber syndrome 1)"
                /note="match: ESTs: AA035463.1 AA134762.1 AA551884.1
                AA603096.1 AA835532.1 AA947591.1 AA977727.1 AI018511.1
                AI052368.1 AI075680.1 AI128442.1 AI200613.1 AI240955.1
                AI313158.1 AI338662.1 AI339819.1 AI346298.1 AI346306.1
                AI346673.1 AI346901.1 AI347971.1 AI470416.1 AI624171.1
                AI808129.1 AI818120.1 AI819006.1 AU134585.1 AU138547.1
                AU139633.1 AU141270.1 AU141705.1 AW665221.1 AW772048.1
                BE463499.1 BE539042.1 BE732652.1 BE901655.1 BF108743.1
                BF203452.1 BF838921.1 BG002252.1 BG012266.1 BG212420.1
                BG571008.1 BG676878.1 BG753449.1 BG772642.1 BG823461.1
                BG826890.1 BI113854.1 BI193718.1 BI259143.1 BI460949.1
                BI521323.1 BI523122.1 BI559359.1 BI559698.1 BI758470.1
                BI771957.1 BI819026.1 BI833372.1 BI838548.1 BI838638.1
                BI860395.1 BM150781.1 BM545942.1 BM665441.1 BM708962.1
                BM725049.1 BM903634.1 BM918321.1 BM970472.1 BM980518.1
                BQ015551.1 BQ019502.1 BQ019802.1 BQ223499.1 BQ668987.1
                BQ670584.1 BQ672050.1 BQ716736.1 BQ883448.1 BQ893639.1
                BQ925406.1 BQ926230.1 BQ927484.1 BQ936939.1 BQ947960.1
                BQ959748.1 BQ962328.1 BU180470.1 BU630414.1 BU726343.1
                BU731798.1 BU741952.1 BU790380.1 CA308834.1 CA309960.1
                H05168.1 H16457.1 H20424.1 H26536.1 H41805.1 T71758.1
                T99132.1 W37947.1
                match: cDNAs: BC014271.1 BC020391.1 BC029080.1 J05481.1
                S69407.1 X77952.1 Z23142.1"
     gene       join(complement(27622. .27969),complement(16427. .16578),
                complement(3020. .3160),
                complement(AL162586.26:78485. .78647),
                complement(AL162586.26:77670. .77835),
                complement(AL162586.26:77206. .77332),
                complement(AL162586.26:76775. .76949),
                complement(AL162586.26:76279. .76421),
                complement(AL162586.26:71875. .72012),
                complement(AL162586.26:71597. .71635),
                complement(AL162586.26:70691. .70807),
                complement(AL162586.26:70095. .70352),
                complement(AL162586.26:69124. .69178),
                complement(AL162586.26:66992. .68028))
                /gene="ENG"
                /locus_tag="RP11-228B15.2-001"
     mRNA       join(complement(27622. .27969),complement(16427. .16578),
                complement(3020. .3160),
                complement(AL162586.26:78485. .78647),
                complement(AL162586.26:77670. .77835),
                complement(AL162586.26:77206. .77332),
                complement(AL162586.26:76775. .76949),
                complement(AL162586.26:76279. .76421),
                complement(AL162586.26:71875. .72012),
                complement(AL162586.26:71597. .71635),
                complement(AL162586.26:70691. .70807),
                complement(AL162586.26:70095. .70352),
                complement(AL162586.26:69124. .69178),
                complement(AL162586.26:66992. .68028))
                /gene="ENG"
                /locus_tag="RP11-228B15.2-001"
                /product="endoglin (Osler-Rendu-Weber syndrome 1)"
                /note="match: ESTs: AW581677.1 BE735050.1 BE909609.1
                BF092020.1 BF345575.1 BG677489.1 BI772291.1 BI861060.1
                BM558848.1 BM560132.1 BM915504.1 BM916319.1 BM921548.1
                BM921566.1 BQ187083.1 BQ189035.1 CA393306.1 H20423.1
                H41845.1 R63870.1
                match: cDNAs: X72012.1"
      CDS       join(complement(27622. .27688),complement(16427. .16578),
                complement(3020. .3160),
                complement(AL162586.26:78485. .78647),
                complement(AL162586.26:77670. .77835),
                complement(AL162586.26:77206. .77332),
                complement(AL162586.26:76775. .76949),
                complement(AL162586.26:76279. .76421),
                complement(AL162586.26:71875. .72012),
                complement(AL162586.26:71597. .71635),
                complement(AL162586.26:70691. .70807),
                complement(AL162586.26:70095. .70352),
                complement(AL162586.26:69124. .69178),
                complement(AL162586.26:67918. .68028),
                complement(AL162586.26:67657. .67781))
                /gene="ENG"
                /locus_tag="RP11-228B15.2-002"
                /standard_name="OTTHUMP00000022222"
                /note="match: proteins: AAH29080 P17813 P37176 Q63961
                Q96CG0"
                /codon_start=1
                /product="endoglin (Osler-Rendu-Weber syndrome 1)"
                /protein_id="CAI12604.1"
                /db_xref="GI:55958253"
                /db_xref="InterPro:IPR001507"
                /translation="MDRGTLPLAVALLLASCSLSPTSLAETVHCDLQPVGPERGEVTY
                TTSQVSKGCVAQAPNAILEVHVLFLEFPTGPSQLELTLQASKQNGTWPREVLLVLSVN
                SSVFLHLQALGIPLHLAYNSSLVTFQEPPGVNTTELPSFPKTQILEWAAERGPITSAA
                ELNDPQSILLRLGQAQGSLSFCMLEASQDMGRTLEWRPRTPALVRGCHLEGVAGHKEA
                HILRVLPGHSAGPRTVTVKVELSCAPGDLDAVLILQGPPYVSWLIDANHNMQIWTTGE
                YSFKIFPEKNIRGFKLPDTPQGLLGEARMLNASIVASFVELPLASIVSLHASSCGGRL
                QTSPAPIQTTPPKDTCSPELLMSLIQTKCADDAMTLVLKKELVAHLKCTITGLTFWDP
                SCEAEDRGDKFVLRSAYSSCGMQVSASMISNEAVVNILSSSSPQRKKVHCLNMDSLSF
                QLGLYLSPHFLQASNTIEPGQQSFVQVRVSPSVSEFLLQLDSCHLDLGPEGGTVELIQ
                GRAAKGNCVSLLSPSPEGDPRFSFLLHFYTVPIPKTGTLSCTVALRPKTGSQDQEVHR
                TVFMRLNIISPDLSGCTSKGLVLPAVLGITFGAFLIGALLTAALWYIYSHTRSPSKRE
                PVVAVAAPASSESSSTNHSIGSTQSTPCSTSSMA"
      CDS       join(complement(27622. .27688),complement(16427. .16578),
                complement(3020. .3160),
                complement(AL162586.26:78485. .78647),
                complement(AL162586.26:77670. .77835),
```

```
  Query Match             62.1%;  Score 18;  DB 8;  Length 164201;
  Best Local Similarity   100.0%;  Pred. No. 2.4;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAGGC 18
              ||||||||||||||||||
Db     102656 CCTCTCTGGAGCTCAGGC 102639


RESULT 12
BX677664
LOCUS       BX677664              170975 bp    DNA     linear   VRT 17-SEP-2004
DEFINITION  Zebrafish DNA sequence from clone DKEYP-122B11 in linkage group 2,
            complete sequence.
ACCESSION   BX677664
VERSION     BX677664.9  GI:52313372
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 170975)
  AUTHORS   Garnett,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-SEP-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Sep 17, 2004 this sequence version replaced gi:51965276.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml
            DKEYP-122B11 is from a Zebrafish BAC library
            VECTOR: pIndigoBAC-5.
FEATURES             Location/Qualifiers
     source          1. .170975
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEYP-122B11"
                     /clone_lib="DanioKeypilot"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 5;  Length 170975;
  Best Local Similarity   100.0%;  Pred. No. 2.4;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 CTCTGGAGCTCAGGCATG 21
              ||||||||||||||||||
Db      71544 CTCTGGAGCTCAGGCATG 71561


RESULT 13
AC155304
LOCUS       AC155304              174243 bp    DNA     linear   HTG 16-JUN-2005
DEFINITION  Mus musculus chromosome 8 clone RP24-212I6, WORKING DRAFT SEQUENCE,
            13 unordered pieces.
ACCESSION   AC155304
VERSION     AC155304.2  GI:67078734
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 174243)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Mus musculus clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 174243)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-2005) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 174243)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUN-2005) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT     On Jun 9, 2005 this sequence version replaced gi:57790222.

            ---------------- Genome Center ----------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            -------------- Project Information -------------------
            Center project name: M_BB0212I06
            ------------------------------------------------------
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 168191 bases at least Q40
            Consensus quality: 169581 bases at least Q30
            Consensus quality: 170484 bases at least Q20
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 13 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1347: contig of 1347 bp in length
            *     1348     1447: gap of unknown length
            *     1448     2945: contig of 1498 bp in length
            *     2946     3045: gap of unknown length
            *     3046     4174: contig of 1129 bp in length
            *     4175     4274: gap of unknown length
            *     4275     5589: contig of 1315 bp in length
            *     5590     5689: gap of unknown length
            *     5690     7104: contig of 1415 bp in length
            *     7105     7204: gap of unknown length
            *     7205     8437: contig of 1233 bp in length
            *     8438     8537: gap of unknown length
            *     8538     9935: contig of 1398 bp in length
            *     9936    10035: gap of unknown length
            *    10036    12606: contig of 2571 bp in length
            *    12607    12706: gap of unknown length
            *    12707    14944: contig of 2238 bp in length
            *    14945    15044: gap of unknown length
            *    15045    21934: contig of 6890 bp in length
            *    21935    22034: gap of unknown length
            *    22035    39543: contig of 17509 bp in length
            *    39544    39643: gap of unknown length
            *    39644    93493: contig of 53850 bp in length
            *    93494    93593: gap of unknown length
            *    93594   174243: contig of 80650 bp in length.
FEATURES             Location/Qualifiers
     source          1. .174243
                     /organism="Mus musculus"
```

```
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="8"
                     /clone="RP24-212I6"
     misc_feature    1. .1347
                     /note="assembly_name:Contig23"
     gap             1348. .1447
                     /estimated_length=unknown
     misc_feature    1448. .2945
                     /note="assembly_name:Contig25"
     gap             2946. .3045
                     /estimated_length=unknown
     misc_feature    3046. .4174
                     /note="assembly_name:Contig39"
     gap             4175. .4274
                     /estimated_length=unknown
     misc_feature    4275. .5589
                     /note="assembly_name:Contig44"
     gap             5590. .5689
                     /estimated_length=unknown
     misc_feature    5690. .7104
                     /note="assembly_name:Contig46"
     gap             7105. .7204
                     /estimated_length=unknown
     misc_feature    7205. .8437
                     /note="assembly_name:Contig51"
     gap             8438. .8537
                     /estimated_length=unknown
     misc_feature    8538. .9935
                     /note="assembly_name:Contig54"
     gap             9936. .10035
                     /estimated_length=unknown
     misc_feature    10036. .12606
                     /note="assembly_name:Contig55"
     gap             12607. .12706
                     /estimated_length=unknown
     misc_feature    12707. .14944
                     /note="assembly_name:Contig56"
     gap             14945. .15044
                     /estimated_length=unknown
     misc_feature    15045. .21934
                     /note="assembly_name:Contig58"
     gap             21935. .22034
                     /estimated_length=unknown
     misc_feature    22035. .39543
                     /note="assembly_name:Contig59"
     gap             39544. .39643
                     /estimated_length=unknown
     misc_feature    39644. .93493
                     /note="assembly_name:Contig60"
     gap             93494. .93593
                     /estimated_length=unknown
     misc_feature    93594. .174243
                     /note="assembly_name:Contig61"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 14;  Length 174243;
  Best Local Similarity   100.0%;  Pred. No. 2.4;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAGGC 18
              ||||||||||||||||||
Db      38575 CCTCTCTGGAGCTCAGGC 38592


RESULT 14
AC080169
LOCUS       AC080169              182271 bp    DNA     linear   HTG 28-SEP-2000
DEFINITION  Mus musculus clone RP23-347L24, WORKING DRAFT SEQUENCE, 25
            unordered pieces.
ACCESSION   AC080169
VERSION     AC080169.1  GI:10334869
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidae; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 182271)
  AUTHORS   McCombie,W.R., Baker,J.P., Bahret,A., Bal,H., Balija,V.,
            Dedhia,N.N., de la Bastide,M., Huang,E.N., King,L., Kirchoff,K.A.,
            Miller,B., Nascimento,L.U., O'Shaughnessy,A.L., Preston,R.R.,
            Rodriguez,S., Santos,L., Shah,R.S., Spiegel,L.A., Toth,K., Vil,M.D.
            and Zutavern,T.
  TITLE     Mouse Genomic Sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 182271)
  AUTHORS   McCombie,W.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-SEP-2000) Lita Annenberg Hazen Genome Sequencing
            Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
            Harbor, NY 11724, USA
COMMENT     ------------------- Genome Center
            Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor
            Laboratory
            Center code: CSHL
            Web site: http://www.cshl.org/genseq
            Contact:  mccombie@cshl.org
            ------------------- Project Information
            Center project name: RP23-347L24
            Center clone name: RP23-347L24
            -------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 25 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    24067: contig of 24067 bp in length
            *    24068    24296: gap of unknown length
            *    24297    39626: contig of 15330 bp in length
            *    39627    39855: gap of unknown length
            *    39856    53099: contig of 13244 bp in length
            *    53100    53327: gap of unknown length
            *    53328    65362: contig of 12035 bp in length
            *    65363    65590: gap of unknown length
            *    65591    75189: contig of 9599 bp in length
            *    75190    75417: gap of unknown length
            *    75418    84074: contig of 8657 bp in length
            *    84075    84302: gap of unknown length
            *    84303    92879: contig of 8577 bp in length
            *    92880    93107: gap of unknown length
            *    93108   101426: contig of 8319 bp in length
            *   101427   101654: gap of unknown length
            *   101655   109861: contig of 8207 bp in length
            *   109862   110089: gap of unknown length
            *   110090   118237: contig of 8148 bp in length
            *   118238   118465: gap of unknown length
            *   118466   125095: contig of 6630 bp in length
            *   125096   125323: gap of unknown length
            *   125324   131302: contig of 5979 bp in length
            *   131303   131530: gap of unknown length
            *   131531   137273: contig of 5743 bp in length
            *   137274   137501: gap of unknown length
            *   137502   141714: contig of 4213 bp in length
            *   141715   141942: gap of unknown length
            *   141943   145917: contig of 3975 bp in length
            *   145918   146145: gap of unknown length
            *   146146   150067: contig of 3922 bp in length
            *   150068   150295: gap of unknown length
            *   150296   154163: contig of 3868 bp in length
            *   154164   154391: gap of unknown length
            *   154392   158095: contig of 3704 bp in length
            *   158096   158323: gap of unknown length
```

```
          *   158324   162022: contig of 3699 bp in length
          *   162023   162250: gap of unknown length
          *   162251   165898: contig of 3648 bp in length
          *   165899   166126: gap of unknown length
          *   166127   169548: contig of 3422 bp in length
          *   169549   169776: gap of unknown length
          *   169777   173131: contig of 3355 bp in length
          *   173132   173359: gap of unknown length
          *   173360   176306: contig of 2947 bp in length
          *   176307   176534: gap of unknown length
          *   176535   179462: contig of 2928 bp in length
          *   179463   179690: gap of unknown length
          *   179691   182271: contig of 2581 bp in length.
FEATURES             Location/Qualifiers
     source          1. .182271
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="RP23-347L24"
                     /clone_lib="RPCI-23"
     gap             24068. .24296
                     /estimated_length=unknown
     gap             39627. .39855
                     /estimated_length=unknown
     gap             53100. .53327
                     /estimated_length=unknown
     gap             65363. .65590
                     /estimated_length=unknown
     gap             75190. .75417
                     /estimated_length=unknown
     gap             84075. .84302
                     /estimated_length=unknown
     gap             92880. .93107
                     /estimated_length=unknown
     gap             101427. .101654
                     /estimated_length=unknown
     gap             109862. .110089
                     /estimated_length=unknown
     gap             118238. .118465
                     /estimated_length=unknown
     gap             125096. .125323
                     /estimated_length=unknown
     gap             131303. .131530
                     /estimated_length=unknown
     gap             137274. .137501
                     /estimated_length=unknown
     gap             141715. .141942
                     /estimated_length=unknown
     gap             145918. .146145
                     /estimated_length=unknown
     gap             150068. .150295
                     /estimated_length=unknown
     gap             154164. .154391
                     /estimated_length=unknown
     gap             158096. .158323
                     /estimated_length=unknown
     gap             162023. .162250
                     /estimated_length=unknown
     gap             165899. .166126
                     /estimated_length=unknown
     gap             169549. .169776
                     /estimated_length=unknown
     gap             173132. .173359
                     /estimated_length=unknown
     gap             176307. .176534
                     /estimated_length=unknown
     gap             179463. .179690
                     /estimated_length=unknown
ORIGIN

  Query Match           62.1%;  Score 18;  DB 14;  Length 182271;
  Best Local Similarity  100.0%;  Pred. No. 2.4;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAGGC 18
              ||||||||||||||||||
Db      80525 CCTCTCTGGAGCTCAGGC 80542


RESULT 15
AC119725/c
LOCUS       AC119725              200288 bp    DNA     linear   HTG 27-MAR-2003
DEFINITION  Homo sapiens chromosome 3 clone RP11-759D18, WORKING DRAFT
            SEQUENCE, 36 unordered pieces.
ACCESSION   AC119725
VERSION     AC119725.2  GI:29293997
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 200288)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 200288)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 200288)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
```

```
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Mar 27, 2003 this sequence version replaced gi:20376806.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Drafting Center Code: BCM
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: HCIR
                Center clone name: RP11-759D18
            -------------- Summary Statistics
                Sequencing vector: Plasmid;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 169013 bases at least Q40
                Consensus quality: 174569 bases at least Q30
                Consensus quality: 178019 bases at least Q20
                Estimated insert size: 177488; sum-of-contigs estimation
                Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 36 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1266: contig of 1266 bp in length
            *     1267     1366: gap of unknown length
            *     1367     2448: contig of 1082 bp in length
            *     2449     2548: gap of unknown length
            *     2549     4153: contig of 1605 bp in length
            *     4154     4253: gap of unknown length
            *     4254     5318: contig of 1065 bp in length
            *     5319     5418: gap of unknown length
            *     5419     6987: contig of 1569 bp in length
            *     6988     7087: gap of unknown length
            *     7088     8369: contig of 1282 bp in length
            *     8370     8469: gap of unknown length
            *     8470     9488: contig of 1019 bp in length
            *     9489     9588: gap of unknown length
            *     9589    11671: contig of 2083 bp in length
            *    11672    11771: gap of unknown length
            *    11772    13295: contig of 1524 bp in length
            *    13296    13395: gap of unknown length
            *    13396    14422: contig of 1027 bp in length
            *    14423    14522: gap of unknown length
            *    14523    15982: contig of 1460 bp in length
            *    15983    16082: gap of unknown length
            *    16083    18344: contig of 2262 bp in length
            *    18345    18444: gap of unknown length
            *    18445    20093: contig of 1649 bp in length
            *    20094    20193: gap of unknown length
            *    20194    21347: contig of 1154 bp in length
            *    21348    21447: gap of unknown length
            *    21448    24953: contig of 3506 bp in length
            *    24954    25053: gap of unknown length
            *    25054    27672: contig of 2619 bp in length
            *    27673    27772: gap of unknown length
            *    27773    29866: contig of 2094 bp in length
            *    29867    29966: gap of unknown length
            *    29967    31985: contig of 2019 bp in length
            *    31986    32085: gap of unknown length
            *    32086    34334: contig of 2249 bp in length
            *    34335    34434: gap of unknown length
            *    34435    37766: contig of 3332 bp in length
            *    37767    37866: gap of unknown length
            *    37867    40539: contig of 2673 bp in length
            *    40540    40639: gap of unknown length
            *    40640    44375: contig of 3736 bp in length
            *    44376    44475: gap of unknown length
            *    44476    50326: contig of 5851 bp in length
            *    50327    50426: gap of unknown length
            *    50427    53739: contig of 3313 bp in length
            *    53740    53839: gap of unknown length
            *    53840    61829: contig of 7990 bp in length
            *    61830    61929: gap of unknown length
            *    61930    68906: contig of 6977 bp in length
            *    68907    69006: gap of unknown length
            *    69007    78680: contig of 9674 bp in length
            *    78681    78780: gap of unknown length
            *    78781    89431: contig of 10651 bp in length
            *    89432    89531: gap of unknown length
            *    89532   100730: contig of 11199 bp in length
            *   100731   100830: gap of unknown length
            *   100831   112190: contig of 11360 bp in length
            *   112191   112290: gap of unknown length
            *   112291   122606: contig of 10316 bp in length
            *   122607   122706: gap of unknown length
            *   122707   135164: contig of 12458 bp in length
            *   135165   135264: gap of unknown length
            *   135265   147115: contig of 11851 bp in length
            *   147116   147215: gap of unknown length
            *   147216   159369: contig of 12154 bp in length
            *   159370   159469: gap of unknown length
            *   159470   173326: contig of 13857 bp in length
            *   173327   173426: gap of unknown length
            *   173427   200288: contig of 26862 bp in length.
FEATURES             Location/Qualifiers
     source          1. .200288
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /clone="RP11-759D18"
     gap             1267. .1366
                     /estimated_length=unknown
     gap             2449. .2548
                     /estimated_length=unknown
     gap             4154. .4253
                     /estimated_length=unknown
     gap             5319. .5418
                     /estimated_length=unknown
     gap             6988. .7087
                     /estimated_length=unknown
     gap             8370. .8469
                     /estimated_length=unknown
     gap             9489. .9588
                     /estimated_length=unknown
     gap             11672. .11771
                     /estimated_length=unknown
     gap             13296. .13395
                     /estimated_length=unknown
     gap             14423. .14522
                     /estimated_length=unknown
     gap             15983. .16082
                     /estimated_length=unknown
     gap             18345. .18444
                     /estimated_length=unknown
     gap             20094. .20193
                     /estimated_length=unknown

  Query Match             62.1%;  Score 18;  DB 14;  Length 200288;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAGGC 18
              ||||||||||||||||||
Db     125376 CCTCTCTGGAGCTCAGGC 125359


RESULT 16
```

```
AL954146
LOCUS       AL954146              207669 bp    DNA     linear   VRT 15-AUG-2003
DEFINITION  Zebrafish DNA sequence from clone CH211-213I6 in linkage group 17,
            complete sequence.
ACCESSION   AL954146
VERSION     AL954146.8  GI:33694361
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 207669)
  AUTHORS   Lloyd,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-AUG-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 15, 2003 this sequence version replaced gi:28625337.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep CH211-213I6 is
            from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1
            Clone-derived Zebrafish pUC subclones occasionally display
            inconsistency over the length of mononucleotide A/T runs and
            conserved TA repeats. Where this is found the longest good quality
            representation will be submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml.
FEATURES             Location/Qualifiers
     source          1. .207669
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-213I6"
                     /clone_lib="CHORI-211"
ORIGIN

  Query Match           62.1%;  Score 18;  DB 5;  Length 207669;
  Best Local Similarity  100.0%;  Pred. No. 2.3;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 CTCTGGAGCTCAGGCATG 21
              ||||||||||||||||||
Db      34436 CTCTGGAGCTCAGGCATG 34453
```

```
RESULT 17
AC126205
LOCUS       AC126205              221370 bp    DNA     linear   HTG 20-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-164I5, WORKING DRAFT SEQUENCE, 2
            unordered pieces.
ACCESSION   AC126205
VERSION     AC126205.4  GI:25138634
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 221370)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,P.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 221370)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 221370)
  AUTHORS   Rat Genome Sequencing Consortium.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 20, 2002 this sequence version replaced gi:22855997.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GZGZ
                Center clone name: CH230-164I5
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 211417 bases at least Q40
                Consensus quality: 213594 bases at least Q30
                Consensus quality: 214957 bases at least Q20
                Estimated insert size: 217908; sum-of-contigs estimation
                Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1   219823: contig of 219823 bp in length
            *  219824   219923: gap of unknown length
            *  219924   221370: contig of 1447 bp in length.
FEATURES             Location/Qualifiers
     source          1. .221370
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-164I5"
     misc_feature    1. .1678
                     /note="wgs_contig"
     misc_feature    1905. .4301
                     /note="wgs_contig"
     gap             219824. .219923
                     /estimated_length=unknown
ORIGIN

  Query Match             62.1%;  Score 18;  DB 14;  Length 221370;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 CTGGAGCTCAGGCATGAG 23
              ||||||||||||||||||
Db     104134 CTGGAGCTCAGGCATGAG 104151


RESULT 18
AC107010
LOCUS       AC107010               248701 bp    DNA     linear   HTG 10-OCT-2002
DEFINITION  Rattus norvegicus clone CH230-195C7, WORKING DRAFT SEQUENCE.
ACCESSION   AC107010
VERSION     AC107010.4  GI:22856953
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 248701)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 248701)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 248701)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-OCT-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Sep 14, 2002 this sequence version replaced gi:21737591.
```

```
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GPKM
                Center clone name: CH230-195C7
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 220459 bases at least Q40
                Consensus quality: 222411 bases at least Q30
                Consensus quality: 223440 bases at least Q20
                Estimated insert size: 224528; sum-of-contigs estimation
                Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1   248701: contig of 248701 bp in length.
FEATURES             Location/Qualifiers
     source          1. .248701
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-195C7"
     misc_feature    1. .1187
                     /note="wgs_contig"
     misc_feature    14577. .16609
                     /note="wgs_contig"
     misc_feature    247661. .248701
                     /note="wgs_contig"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 14;  Length 248701;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 CTGGAGCTCAGGCATGAG 23
              ||||||||||||||||||
Db     221754 CTGGAGCTCAGGCATGAG 221771


RESULT 19
BV323421
LOCUS       BV323421                650 bp    DNA     linear   STS 26-JAN-2005
DEFINITION  S236P6539RG8.T0 AlaskanMalamute Canis familiaris STS genomic,
            sequence tagged site.
ACCESSION   BV323421
VERSION     BV323421.1  GI:57521913
KEYWORDS    STS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
            Canis.
  REFERENCE 1  (bases 1 to 650)
    AUTHORS   Lindblad-Toh,K.
    TITLE     The genome sequence of Canis familiaris
    JOURNAL   Unpublished (2004)
  COMMENT

            Contact: Kerstin Lindblad-Toh
            Whitehead Institute for Biomedical Research, Center for Genome
            Research
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172521477
            Fax: 6172580903
            Email: kersli@genome.wi.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 650
            Protocol:
            WGS-discovery (WGS):
            Paired-end low-coverage whole genome shotgun reads were generated
            from 9 breeds
            (German Shepherd, Rottweiler, Bedlington Terrier, Beagle, Labrador
            Retriever, English
            Shepherd, Italian Greyhound, Alaskan Malamute and the Portuguese
            Water Dog -100,000 each)
            and five other canids (Chinese, Alaskan, Indian and Spanish Gray
            Wolf as well as the
            Californian Coyote).
            The WGS reads were placed uniquely on the CanFam1.0 boxer assembly
            and SNP detection was
            carried out by SSAHA-SNP. 863872 reads were annotated as STSs and
            485941 SNPs were
            annotated with alleles from the boxer and the breed or canid from
            which the particular
            read came. The validation rate for these SNPs was estimated at
            approximately 98%.
            WGA-discovery (WGA) of Boxer/Poodle SNPs:
            A second set of SNPs was generated using a similar methodology
            except that the contigs
            from the 1.5x poodle assembly (Kirkness 2003) were used instead of
            WGS reads. Since this
            sequence lacked base quality scores, arbitrary quality scores of
            phred 40 were assigned
            before the poddle seqeunce was placed uniquely on the CanFam1.0
            boxer assembly and SNP
            detection was carried out by SSAHA-SNP. 1637780 SNPs were annotated
            with alleles from the
            boxer and the poodle. The validation rate for these SNPs was
            estimated at approximately TBD%.
            Internal-WGA-discovery (I-WGA):
            A third set of SNPs were discovered by comparing reads in the WGA
            assembly. SNPs were
            defined as mismatch positions that had a base quality of >= 30 on
            both reads in a region
            that aligned without gaps, and with at most one additional mismatch
            in the ten flanking
            bases. For each allele, at least one additional read had to confirm
            it. 731476 SNPs were
            annotated with alleles between the two boxer alleles. The
            validation rate for these SNPs
            was estimated at approximately TBD%.
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="AlaskanMalamute"
                     /db_xref="taxon:9615"
                     /map="+ 9 22-606 25058368-25057784"
                     /clone_lib="AlaskanMalamute"
     STS             <1. .>650
ORIGIN
```

```
  Query Match             58.6%;  Score 17;  DB 10;  Length 650;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db        220 TCAGGCATGAGCCAGCA 236


RESULT 20
F272846S27
LOCUS       F272846S27              1006 bp    DNA     linear   PRI 14-MAR-2001
DEFINITION  Homo sapiens Fanconi anemia complementation group D2 protein
            (FANCD2) gene, exon 39.
ACCESSION   AF273247
VERSION     AF273247.1  GI:13324517
KEYWORDS    .
SEGMENT     27 of 31
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1006)
  AUTHORS   Timmers,C., Taniguchi,T., Hejna,J., Reifsteck,C., Lucas,L.,
            Bruun,D., Thayer,M., Cox,B., Olson,S., D'Andrea,A.D., Moses,R. and
            Grompe,M.
  TITLE     Positional cloning of a novel Fanconi anemia gene, FANCD2
  JOURNAL   Mol. Cell 7 (2), 241-248 (2001)
   PUBMED   11239453
REFERENCE   2  (bases 1 to 1006)
  AUTHORS   Timmers,C.D. and Grompe,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-2000) Molecular & Medical Genetics, Oregon Health
            Sciences University, 3181 S.W. Sam Jackson Park Rd. Mail Code L103,
            Portland, OR 97201, USA
FEATURES             Location/Qualifiers
     source          1. .1006
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /map="3p26"
     exon            572. .610
                     /gene="FANCD2"
                     /number=39
ORIGIN

  Query Match             58.6%;  Score 17;  DB 8;  Length 1006;
  Best Local Similarity   100.0%;  Pred. No. 21;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db        253 TCAGGCATGAGCCAGCA 269


RESULT 21
AC105975/c
LOCUS       AC105975               68409 bp    DNA     linear   HTG 11-JAN-2002
DEFINITION  Mus musculus clone RP24-409G7, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC105975
VERSION     AC105975.1  GI:18129475
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 68409)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus, clone RP24-409G7
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 68409)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L20174
                Center clone name: 409_G_7
            --------------.
            * NOTE: This record contains 90 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      650: contig of 650 bp in length
            *      651      750: gap of 100 bp
            *      751     1402: contig of 652 bp in length
            *     1403     1502: gap of 100 bp
            *     1503     2155: contig of 653 bp in length
            *     2156     2255: gap of 100 bp
            *     2256     2922: contig of 667 bp in length
            *     2923     3022: gap of 100 bp
            *     3023     3685: contig of 663 bp in length
            *     3686     3785: gap of 100 bp
            *     3786     4452: contig of 667 bp in length
            *     4453     4552: gap of 100 bp
            *     4553     5223: contig of 671 bp in length
            *     5224     5323: gap of 100 bp
            *     5324     6049: contig of 726 bp in length
            *     6050     6149: gap of 100 bp
            *     6150     6806: contig of 657 bp in length
            *     6807     6906: gap of 100 bp
            *     6907     7569: contig of 663 bp in length
            *     7570     7669: gap of 100 bp
            *     7670     8322: contig of 653 bp in length
            *     8323     8422: gap of 100 bp
            *     8423     9160: contig of 738 bp in length
            *     9161     9260: gap of 100 bp
```

```
*     9261     9910: contig of 650 bp in length
*     9911    10010: gap of 100 bp
*    10011    10653: contig of 643 bp in length
*    10654    10753: gap of 100 bp
*    10754    11415: contig of 662 bp in length
*    11416    11515: gap of 100 bp
*    11516    12186: contig of 671 bp in length
*    12187    12286: gap of 100 bp
*    12287    12952: contig of 666 bp in length
*    12953    13052: gap of 100 bp
*    13053    13718: contig of 666 bp in length
*    13719    13818: gap of 100 bp
*    13819    14457: contig of 639 bp in length
*    14458    14557: gap of 100 bp
*    14558    15218: contig of 661 bp in length
*    15219    15318: gap of 100 bp
*    15319    15977: contig of 659 bp in length
*    15978    16077: gap of 100 bp
*    16078    16726: contig of 649 bp in length
*    16727    16826: gap of 100 bp
*    16827    17478: contig of 652 bp in length
*    17479    17578: gap of 100 bp
*    17579    18234: contig of 656 bp in length
*    18235    18334: gap of 100 bp
*    18335    18985: contig of 651 bp in length
*    18986    19085: gap of 100 bp
*    19086    19714: contig of 629 bp in length
*    19715    19814: gap of 100 bp
*    19815    20476: contig of 662 bp in length
*    20477    20576: gap of 100 bp
*    20577    21240: contig of 664 bp in length
*    21241    21340: gap of 100 bp
*    21341    22013: contig of 673 bp in length
*    22014    22113: gap of 100 bp
*    22114    22778: contig of 665 bp in length
*    22779    22878: gap of 100 bp
*    22879    23543: contig of 665 bp in length
*    23544    23643: gap of 100 bp
*    23644    24385: contig of 742 bp in length
*    24386    24485: gap of 100 bp
*    24486    25146: contig of 661 bp in length
*    25147    25246: gap of 100 bp
*    25247    25899: contig of 653 bp in length
*    25900    25999: gap of 100 bp
*    26000    26649: contig of 650 bp in length
*    26650    26749: gap of 100 bp
*    26750    27387: contig of 638 bp in length
*    27388    27487: gap of 100 bp
*    27488    28156: contig of 669 bp in length
*    28157    28256: gap of 100 bp
*    28257    28915: contig of 659 bp in length
*    28916    29015: gap of 100 bp
*    29016    29654: contig of 639 bp in length
*    29655    29754: gap of 100 bp
*    29755    30416: contig of 662 bp in length
*    30417    30516: gap of 100 bp
*    30517    31181: contig of 665 bp in length
*    31182    31281: gap of 100 bp
*    31282    31938: contig of 657 bp in length
*    31939    32038: gap of 100 bp
*    32039    32700: contig of 662 bp in length
*    32701    32800: gap of 100 bp
*    32801    33465: contig of 665 bp in length
*    33466    33565: gap of 100 bp
*    33566    34200: contig of 635 bp in length
*    34201    34300: gap of 100 bp
*    34301    34945: contig of 645 bp in length
*    34946    35045: gap of 100 bp
*    35046    35704: contig of 659 bp in length
*    35705    35804: gap of 100 bp
*    35805    36454: contig of 650 bp in length
*    36455    36554: gap of 100 bp
*    36555    37199: contig of 645 bp in length
*    37200    37299: gap of 100 bp
*    37300    37958: contig of 659 bp in length
*    37959    38058: gap of 100 bp
*    38059    38713: contig of 655 bp in length
*    38714    38813: gap of 100 bp
*    38814    39472: contig of 659 bp in length
*    39473    39572: gap of 100 bp
*    39573    40228: contig of 656 bp in length
*    40229    40328: gap of 100 bp
*    40329    41000: contig of 672 bp in length
*    41001    41100: gap of 100 bp
*    41101    41757: contig of 657 bp in length
*    41758    41857: gap of 100 bp
*    41858    42511: contig of 654 bp in length
*    42512    42611: gap of 100 bp
*    42612    43270: contig of 659 bp in length
*    43271    43370: gap of 100 bp
*    43371    44165: contig of 795 bp in length
*    44166    44265: gap of 100 bp
*    44266    44925: contig of 660 bp in length
*    44926    45025: gap of 100 bp
*    45026    45688: contig of 663 bp in length
*    45689    45788: gap of 100 bp
*    45789    46433: contig of 645 bp in length
*    46434    46533: gap of 100 bp
*    46534    47202: contig of 669 bp in length
*    47203    47302: gap of 100 bp
*    47303    47956: contig of 654 bp in length
*    47957    48056: gap of 100 bp
*    48057    48726: contig of 670 bp in length
*    48727    48826: gap of 100 bp
*    48827    49374: contig of 548 bp in length
*    49375    49474: gap of 100 bp
*    49475    50138: contig of 664 bp in length
*    50139    50238: gap of 100 bp
*    50239    50889: contig of 651 bp in length
*    50890    50989: gap of 100 bp
*    50990    51638: contig of 649 bp in length
*    51639    51738: gap of 100 bp
*    51739    52401: contig of 663 bp in length

  Query Match             58.6%;  Score 17;  DB 14;  Length 68409;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         10 AGCTCAGGCATGAGCCA 26
              |||||||||||||||||
Db      27615 AGCTCAGGCATGAGCCA 27599


RESULT 22
AC116403/c
LOCUS       AC116403               72300 bp    DNA     linear   HTG 19-SEP-2002
DEFINITION  Mus musculus clone RP23-143I4, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC116403
VERSION     AC116403.3  GI:23196455
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 72300)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus, clone RP23-143I4
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 72300)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
```

```
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 72300)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-SEP-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 19, 2002 this sequence version replaced gi:20800274.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L23427
                Center clone name: 143_I_4
            --------------.
            * NOTE: This record contains 87 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      734: contig of 734 bp in length
            *      735      834: gap of 100 bp
            *      835     1569: contig of 735 bp in length
            *     1570     1669: gap of 100 bp
            *     1670     2416: contig of 747 bp in length
            *     2417     2516: gap of 100 bp
            *     2517     3264: contig of 748 bp in length
            *     3265     3364: gap of 100 bp
            *     3365     4049: contig of 685 bp in length
            *     4050     4149: gap of 100 bp
            *     4150     4892: contig of 743 bp in length
            *     4893     4992: gap of 100 bp
            *     4993     5701: contig of 709 bp in length
            *     5702     5801: gap of 100 bp
            *     5802     6546: contig of 745 bp in length
            *     6547     6646: gap of 100 bp
            *     6647     7371: contig of 725 bp in length
            *     7372     7471: gap of 100 bp
            *     7472     8203: contig of 732 bp in length
            *     8204     8303: gap of 100 bp
            *     8304     9032: contig of 729 bp in length
            *     9033     9132: gap of 100 bp
            *     9133     9860: contig of 728 bp in length
            *     9861     9960: gap of 100 bp
            *     9961    10690: contig of 730 bp in length
            *    10691    10790: gap of 100 bp
            *    10791    11531: contig of 741 bp in length
            *    11532    11631: gap of 100 bp
            *    11632    12363: contig of 732 bp in length
            *    12364    12463: gap of 100 bp
            *    12464    13209: contig of 746 bp in length
            *    13210    13309: gap of 100 bp
            *    13310    14060: contig of 751 bp in length
            *    14061    14160: gap of 100 bp
            *    14161    14894: contig of 734 bp in length
            *    14895    14994: gap of 100 bp
            *    14995    15694: contig of 700 bp in length
            *    15695    15794: gap of 100 bp
            *    15795    16514: contig of 720 bp in length
            *    16515    16614: gap of 100 bp
            *    16615    17339: contig of 725 bp in length
            *    17340    17439: gap of 100 bp
            *    17440    18181: contig of 742 bp in length
            *    18182    18281: gap of 100 bp
            *    18282    19033: contig of 752 bp in length
            *    19034    19133: gap of 100 bp
            *    19134    19832: contig of 699 bp in length
            *    19833    19932: gap of 100 bp
            *    19933    20675: contig of 743 bp in length
            *    20676    20775: gap of 100 bp
            *    20776    21524: contig of 749 bp in length
            *    21525    21624: gap of 100 bp
            *    21625    22361: contig of 737 bp in length
            *    22362    22461: gap of 100 bp
            *    22462    23165: contig of 704 bp in length
            *    23166    23265: gap of 100 bp
            *    23266    24015: contig of 750 bp in length
            *    24016    24115: gap of 100 bp
            *    24116    24859: contig of 744 bp in length
            *    24860    24959: gap of 100 bp
            *    24960    25686: contig of 727 bp in length
            *    25687    25786: gap of 100 bp
            *    25787    26511: contig of 725 bp in length
            *    26512    26611: gap of 100 bp
            *    26612    27349: contig of 738 bp in length
            *    27350    27449: gap of 100 bp
            *    27450    28184: contig of 735 bp in length
            *    28185    28284: gap of 100 bp
            *    28285    29021: contig of 737 bp in length
            *    29022    29121: gap of 100 bp
            *    29122    29840: contig of 719 bp in length
            *    29841    29940: gap of 100 bp
            *    29941    30659: contig of 719 bp in length
            *    30660    30759: gap of 100 bp
            *    30760    31507: contig of 748 bp in length
            *    31508    31607: gap of 100 bp
            *    31608    32352: contig of 745 bp in length
            *    32353    32452: gap of 100 bp
            *    32453    33199: contig of 747 bp in length
            *    33200    33299: gap of 100 bp
            *    33300    34028: contig of 729 bp in length
```

```
            *    34029    34128: gap of 100 bp
            *    34129    34835: contig of 707 bp in length
            *    34836    34935: gap of 100 bp
            *    34936    35656: contig of 721 bp in length
            *    35657    35756: gap of 100 bp
            *    35757    36489: contig of 733 bp in length
            *    36490    36589: gap of 100 bp
            *    36590    37328: contig of 739 bp in length
            *    37329    37428: gap of 100 bp
            *    37429    38164: contig of 736 bp in length
            *    38165    38264: gap of 100 bp
            *    38265    39003: contig of 739 bp in length
            *    39004    39103: gap of 100 bp
            *    39104    39839: contig of 736 bp in length
            *    39840    39939: gap of 100 bp
            *    39940    40680: contig of 741 bp in length
            *    40681    40780: gap of 100 bp
            *    40781    41533: contig of 753 bp in length
            *    41534    41633: gap of 100 bp
            *    41634    42381: contig of 748 bp in length
            *    42382    42481: gap of 100 bp
            *    42482    43235: contig of 754 bp in length
            *    43236    43335: gap of 100 bp
            *    43336    44052: contig of 717 bp in length
            *    44053    44152: gap of 100 bp
            *    44153    44883: contig of 731 bp in length
            *    44884    44983: gap of 100 bp
            *    44984    45740: contig of 757 bp in length
            *    45741    45840: gap of 100 bp
            *    45841    46584: contig of 744 bp in length
            *    46585    46684: gap of 100 bp
            *    46685    47416: contig of 732 bp in length
            *    47417    47516: gap of 100 bp

  Query Match             58.6%;  Score 17;  DB 14;  Length 72300;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db      49058 TCAGGCATGAGCCAGCA 49042


RESULT 23
BX649443/c
LOCUS       BX649443               79564 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone CTD-2589D3 on chromosome X Contains
            part of the DHRSX gene for dehydrogenase/reductase (SDR family)
            X-linked and a novel gene, complete sequence.
ACCESSION   BX649443
VERSION     BX649443.16  GI:46406612
KEYWORDS    HTG; DHRSX.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 79564)
  AUTHORS   Bird,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 19, 2004 this sequence version replaced gi:46357786.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome X, constructed by the Sanger Centre Chromosome X Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/ChrX
            CTD-2589D3 is from the CalTech genomic sperm BAC library D. VECTOR:
            pBeloBAC11
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .79564
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="CTD-2589D3"
                     /clone_lib="CIT-HSP-D2"
     gene            join(AC079176.15:85615. .85734,AC079176.15:161247. .161354,
                     AC079176.15:177738. .177806,complement(18653. .18754),
                     complement(CR381696.5:31219. .31426),
                     complement(CR381696.5:<7502. .7709))
                     /gene="DHRSX"
                     /locus_tag="RP11-325D5.2-001"
     gene            join(AC079176.15:<85901. .86240,AC079176.15:161247. .161354,
                     AC079176.15:177738. .177806,complement(18653. .18754),
                     complement(CR381696.5:<31249. .31426))
                     /gene="DHRSX"
                     /locus_tag="RP11-325D5.2-002"
     gene            join(AC079176.15:85578. .85734,AC079176.15:161247. .161354,
                     complement(18653. .18754),
                     complement(CR381696.5:31219. .31426),
                     complement(CR381696.5:<7502. .7709))
                     /gene="DHRSX"
                     /locus_tag="RP11-325D5.2-004"
     gene            join(AC079176.15:85626. .85734,AC079176.15:161247. .161354,
                     AC079176.15:177738. .177806,complement(18653. .18754),
                     complement(CR381696.5:31219. .31426),
                     complement(CR381696.5:<7502. .7709))
                     /gene="DHRSX"
                     /locus_tag="RP11-325D5.2-005"
     mRNA            join(AC079176.15:85578. .85734,AC079176.15:161247. .161354,
                     complement(18653. .18754),
                     complement(CR381696.5:31219. .31426),
                     complement(CR381696.5:<7502. .7709))
                     /gene="DHRSX"
                     /locus_tag="RP11-325D5.2-001"
                     /product="dehydrogenase/reductase (SDR family) X-linked"
                     /note="match: ESTs: Em:BG721504.1"
     mRNA            join(AC079176.15:85615. .85734,AC079176.15:161247. .161354,
                     AC079176.15:177738. .177806,complement(18653. .18754),
                     complement(CR381696.5:31219. .31426),
                     complement(CR381696.5:<7502. .7709))
                     /gene="DHRSX"
                     /locus_tag="RP11-325D5.2-001"
                     /product="dehydrogenase/reductase (SDR family) X-linked"
                     /note="match: ESTs: Em:AA236340.1 Em:BI752210.1
                     Em:BM794858.1 Em:BM796973.1 Em:BM839662.1 Em:BQ685027.1
                     Em:BQ686074.1
                     match: cDNAs: BC032340.1 Em:AJ293620.1 Em:AY358849.1
                     Em:BC009269.1 Em:BC019696.2 Em:BC032340.1"
     mRNA            join(AC079176.15:<85901. .86240,AC079176.15:161247. .161354,
                     AC079176.15:177738. .177806,complement(18653. .18754),
                     complement(CR381696.5:<31249. .31426))
                     /gene="DHRSX"
                     /locus_tag="RP11-325D5.2-001"
                     /product="dehydrogenase/reductase (SDR family) X-linked"
                     /note="match: ESTs: CB961696.1"
     CDS             join(AC079176.15:85626. .85734,AC079176.15:161247. .161354,
```

```
                     complement(18653. .18754),
                     complement(CR381696.5:31219. .31426),
                     complement(CR381696.5:<7502. .7709))
                     /gene="DHRSX"
                     /locus_tag="RP11-325D5.2-004"
                     /standard_name="OTTHUMP00000070385"
                     /codon_start=1
                     /product="dehydrogenase/reductase (SDR family) X-linked"
                     /protein_id="CAI95430.1"
                     /db_xref="GI:66348192"
                     /translation="MSPLSAARAALRVYAVGAAVILAQLLRRCRGGFLEPVFPPRPDR
                     VAIVTGGTDGIGYSTAKHLARLGMHVIIVEFLYCDLASMTSIRQFVQKFKMKKIPLHV
                     LINNAGVMMVPQRKTRDGFEEHFGLNYLGHFLLTNLLLDTLKESGSPGHSARVVTVSS
                     ATHYVAELNMDDLQSSACYSPHAAYAQSKLALVLFTYHLQRLLAAEGSHVTANVVDPG
                     VVNTDVYKHVFWATRLAKKLLGWLLFK"
     CDS             join(AC079176.15:85626. .85734,AC079176.15:161247. .161354,
                     AC079176.15:177738. .177806,complement(18653. .18754),
                     complement(CR381696.5:31219. .31426),
                     complement(CR381696.5:<7502. .7709))
                     /gene="DHRSX"
                     /locus_tag="RP11-325D5.2-005"
                     /standard_name="OTTHUMP00000070386"
                     /codon_start=1
                     /product="dehydrogenase/reductase (SDR family) X-linked"
                     /protein_id="CAI95431.1"
                     /db_xref="GI:66348193"
                     /translation="MSPLSAARAALRVYAVGAAVILAQLLRRCRGGFLEPVFPPRPDR
                     VAIVTGGTDGIGYSTAKHLARLGMHVIIAGNNDSKAKQVVSKIKEETLNDKVEFLYCD
                     LASMTSIRQFVQKFKMKKIPLHVLINNAGVMMVPQRKTRDGFEEHFGLNYLGHFLLTN
                     LLLDTLKESGSPGHSARVVTVSSATHYVAELNMDDLQSSACYSPHAAYAQSKLALVLF
                     TYHLQRLLAAEGSHVTANVVDPGVVNTDVYKHVFWATRLAKKLLGWLLFK"
     CDS             join(AC079176.15:<85901. .86240,AC079176.15:161247. .161354,
                     AC079176.15:177738. .177806,complement(18653. .18754),
                     complement(CR381696.5:<31249. .31426))
                     /gene="DHRSX"
                     /locus_tag="RP11-325D5.2-001"
                     /standard_name="OTTHUMP00000070387"
                     /codon_start=1
                     /product="dehydrogenase/reductase (SDR family) X-linked"
                     /protein_id="CAI95432.1"
                     /db_xref="GI:66348194"
                     /translation="VRAGTSPPRAFRPRRLFCAPGPARPGPTLPSPRVWPPRLSRVGL
                     QPSSPRAAAPPRTLTPPGHGRGPGGSHAPRPRHQAALPAACGRPVTARPRQAPEGRSG
                     RPRDLDPYPPQVFPPRPDRVAIVTGGTDGIGYSTAKHLARLGMHVIIAGNNDSKAKQV
                     VSKIKEETLNDKVEFLYCDLASMTSIRQFVQKFKMKKIPLHVLINNAGVMMVPQRKTR
                     DGFEEHFGLNYLGHFLLTNLLLDTLKESGSPGHSARVVTVSSATHYVA"
     gene            complement(join(<61446. .62070,63536. .63561))
                     /locus_tag="RP13-858C7.1-001"
     mRNA            complement(join(<61446. .62070,63536. .63561))
                     /locus_tag="RP13-858C7.1-001"
                     /product="Putative novel transcript"
                     /note="match: ESTs: Em:BG828262.1"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 8;  Length 79564;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAGG 17
              |||||||||||||||||
Db      19509 CCTCTCTGGAGCTCAGG 19493


RESULT 24
AC007999/c
LOCUS       AC007999               80914 bp    DNA     linear   PRI 01-DEC-2001
DEFINITION  Homo sapiens 3q25-26 BAC CTB-177N7 (California Institute of
            Technology BAC Library) complete sequence.
ACCESSION   AC007999
VERSION     AC007999.13  GI:17223146
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 80914)
  AUTHORS   Muzny,D.M., Adams,C., Bailey,M., Barbaria,J., Blankenburg,K.,
            Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Garcia,D.K., Gorrell,J.H.,
            Gorrell,L.L., Guevara,W., Harris,K., He,X., Hernandez,J.,
            Hodgson,A., Hogues,M., Holloway,C., Hosak,H., Jackson,L.E.,
            Jackson,L., Jia,Y., Jones,M., Kelly,S., Kondejewski,N., Kong,Y.,
            Kovar,C., Leal,B., Li,Z., Lichtarge,O., Liu,J., Liu,W., Logan,O.,
            Lozado,R.J., Lu,J., Lucier,R., Martin,R., Martinez,C., McLeod,M.P.,
            Mei,G., Moore,S., Moorish,T., Morgan,M., Morris,S., Nash,S.,
            Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S., Oswal,G., Parish,B.,
            Paxton,S., Payton,B., Perez,L., Pu,L.L., Quiles,M., Reiter,D.,
            Rives,M., Samuel,S., Say,J., Scherer,S., Shah,E., Shen,H.,
            Simon,M., Sparks,A., Stamps,A., Sucgang,R., Tabor,P., Taylor,T.,
            Vasquez,L., Vinson,R., Vo,Q., Wahbah,M., Watlington,S.,
            Weinstock,G., Weinstock,I.R., Williamson,A., Worley,K., Wren,J.,
            Wrensford,G., Yu,W., Zhou,X., Naylor,S.L., Nelson,D. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 80914)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-1999) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 80914)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-FEB-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 80914)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-FEB-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   5  (bases 1 to 80914)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-DEC-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Dec 1, 2001 this sequence version replaced gi:13112129.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

            CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
               STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
               Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
               Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
            continuity across the splice junctions. Sequences that are not
            identical matches are annotated as similar.
```

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

QUALSTAT-REPORT----------

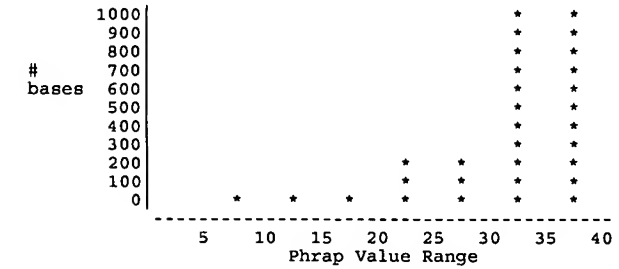
---------- Summary Statistics ----------

Contig length:	82263
Phrap values in estimate:	79909
Average error rate (BCM-Phrap estimate):	0.000185389
Fraction of Phrap values less than 40 :	0.055363
Number of consensus changing edits:	14
Number of N's in consensus :	0

---------- Consensus changing edits ----------

Position	Original+Context	Edited+Context
15239	ggatggggga(n)cggaaaggac	ggatggggga(g)cggaaaggac
15545	agataactat(n)gatattctga	agataactat(t)gatattctga
15642	atggctcatg(n)ctgtaatccc	atggctcatg(c)ctgtaatccc
16126	gctcttgctg(n)ccaggctgga	gctcttgctg(c)ccaggctgga
37467	cgggagaggg(n)gagngnnagg	cgggagaggg(g)gaggggagg
37471	agagggngag(n)gnnaggngga	agaggggggag(g)gggaggggga
37473	agggngagng(n)naggnggagg	agggggaggg(g)gaggggagg
37474	gggngagngn(n)aggnggaggg	gggggagggg(g)agggggaggg
37478	gagngnnagg(n)ggagggggag	gaggggggag(g)ggaggggag
44972	ctccccccctc(n)cccaccccac	ctccccccctc(c)cccaccccac
52504	atatatatat(n)tatatacatg	atatatatat(g)tatatacatg
76365	caagcatttt(n)ntntatttcg	caagcatttt(c)atctatttcg
76366	aagcattttn(n)tntatttcgg	aagcattttc(a)tctatttcgg
76368	gcattttnnt(n)tatttcggtt	gcattttcat(c)tatttcggtt

--------- Distribution of Quality < 40 Bases ----------



Version: 1.01 qxfo.

FEATURES	Location/Qualifiers
source	1. .80914
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="3q25-26"
	/clone="CTB-177N7"
repeat_region	237. .433
	/rpt_family="AluSq/x"
repeat_region	434. .455
	/rpt_family="(TAAAA)n"
gene	join(620. .814,2299. .2485)
	/gene="Unigene cluster containing AA421292 and AA421396"
misc_feature	747. .1253
	/gene="Unigene cluster containing AA421292 and AA421396"
	/note="Region: Unigene cluster containing AW074860 and AA609512"
repeat_region	complement(1799. .1962)
	/rpt_family="AluJb"
repeat_region	complement(1963. .2267)
	/rpt_family="AluSp"
repeat_region	complement(2268. .2413)
	/rpt_family="AluJb"
repeat_region	complement(2515. .2845)
	/rpt_family="AluY"
repeat_region	2946. .3252
	/rpt_family="AluSg"
repeat_region	3706. .4010
	/rpt_family="AluSg"
gene	complement(join(4092. .4200,6573. .6668,8916. .9057, 10117. .10188,12065. .12100))
	/gene="Unigene cluster containing T91383 and AW089025"
repeat_region	complement(4309. .4600)
	/rpt_family="AluSx"
repeat_region	4715. .5006
	/rpt_family="AluSx"
repeat_region	complement(5198. .5474)
	/rpt_family="AluJo"
repeat_region	complement(5616. .5908)
	/rpt_family="AluSg"
repeat_region	complement(5910. .6042)
	/rpt_family="FLAM_C"
STS	5944. .6086
	/standard_name="G26135"
	/db_xref="dbSTS:39988"
misc_feature	6046. .6519
	/note="Region: Unigene cluster containing R94933 and AL046580"
repeat_region	7049. .7233
	/rpt_family="MIR"
repeat_region	7433. .8502
	/rpt_family="MER11C"
misc_feature	9031. .9488
	/note="Region: Unigene cluster containing AA677243 and AL046581"

Query Match 58.6%; Score 17; DB 8; Length 80914;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db      12422 TCAGGCATGAGCCAGCA 12406
```

RESULT 25
HSDJ999L4

LOCUS	HSDJ999L4 100080 bp DNA linear PRI 18-MAY-2005
DEFINITION	Human DNA sequence from clone RP5-999L4 on chromosome 20 Contains the RPL12L3 gene for ribosomal protein L12-like 3 pseudogene and the 5' end of the RIN2 gene for Ras and Rab interactor 2, complete sequence.
ACCESSION	AL132821
VERSION	AL132821.17 GI:7159786
KEYWORDS	HTG; AF-6; FLJ37565; JC265; RAB5 interacting protein 2; RalGDS; RASSF4; RIN2; RPL12L3.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 100080)

```
AUTHORS   Matthews,L.
TITLE     Direct Submission
JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
          Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
          Clone requests: clonerequest@sanger.ac.uk
COMMENT   On Mar 6, 2000 this sequence version replaced gi:7105944.
          The following abbreviations are used to associate primary accession
          numbers given in the feature table with their source databases:
          Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
          on the WORMPEP database can be found at
          http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
          was generated from part of bacterial clone contigs of human
          chromosome 20, constructed by the Sanger Centre Chromosome 20
          Mapping Group.  Further information can be found at
          http://www.sanger.ac.uk/HGP/Chr20
          RP5-999L4 is from the library RPCI-5 constructed by the group of
          Pieter de Jong. For further details see
          http://www.chori.org/bacpac/home.htm
          VECTOR: pCYPAC2
          -------------- Genome Center
          Center: Wellcome Trust Sanger Institute
          Center code: SC
          Web site: http://www.sanger.ac.uk
          Contact: vega@sanger.ac.uk
          --------------
          This sequence was finished as follows unless otherwise noted: all
          regions were either double-stranded or sequenced with an alternate
          chemistry or covered by high quality data (i.e., phred quality >=
          30); an attempt was made to resolve all sequencing problems, such
          as compressions and repeats; all regions were covered by at least
          one subclone; and the assembly was confirmed by restriction digest,
          except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .100080
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIP704L04999"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /clone="RP5-999L4"
                     /clone_lib="RPCI-5"
     misc_feature    1
                     /note="Clone_left_end: RP5-999L4"
     polyA_site      complement(885)
     misc_feature    898
                     /note="Clone_right_end: RP1-122P22"
     polyA_signal    complement(900. .905)
     gene            complement(922. .1422)
                     /locus_tag="RP5-999L4.1-001"
                     /pseudo
     CDS             complement(922. .1422)
                     /locus_tag="RP5-999L4.1-001"
                     /note="match: proteins: O29712 O50003 O60886 O62290 O75000
                     P17079 P23358 P30050 P35979 P50884 Q9ZSL1"
                     /pseudo
                     /codon_start=1
                     /product="ribosomal protein L12 (Rpl12) pseudogene"
     gene            join(64000. .64259,67045. .67137,AL049538.9:12596. .12696,
                     AL049538.9:34112. .34304,AL049538.9:38197. .38308,
                     AL049538.9:42449. .42521,AL049538.9:48335. .48426,
                     AL049538.9:52151. .53284,AL049538.9:67503. .67808,
                     AL049538.9:69665. .69796,AL049538.9:74176. .74339,
                     AL049538.9:78110. .79954)
                     /gene="RIN2"
                     /locus_tag="RP5-1175I6.1-003"
     mRNA            join(64000. .64259,67045. .67137,AL049538.9:12596. .12696,
                     AL049538.9:34112. .34304,AL049538.9:38197. .38308,
                     AL049538.9:42449. .42521,AL049538.9:48335. .48426,
                     AL049538.9:52151. .53284,AL049538.9:67503. .67808,
                     AL049538.9:69665. .69796,AL049538.9:74176. .74339,
                     AL049538.9:78110. .79954)
                     /gene="RIN2"
                     /locus_tag="RP5-1175I6.1-003"
                     /product="Ras and Rab interactor 2"
                     /note="match: ESTs: Em:AI696974.1 Em:AI805157.1
                     Em:AW243662.1 Em:BE938433.1 Em:BM550342.1 Em:BU156049.1
                     Em:CB850876.1 Em:R83223.1 Em:T27862.1
                     match: cDNAs: Em:AB060339.1 Em:AK014548.1 Em:AK040763.1
                     Em:AK094884.1 Em:AL136924.1 Em:BC005529.1 Em:M37190.1"
     CDS             join(64149. .64259,67045. .67137,AL049538.9:12596. .12696,
                     AL049538.9:34112. .34304,AL049538.9:38197. .38308,
                     AL049538.9:42449. .42521,AL049538.9:48335. .48426,
                     AL049538.9:52151. .53284,AL049538.9:67503. .67808,
                     AL049538.9:69665. .69796,AL049538.9:74176. .74339,
                     AL049538.9:78110. .78433)
                     /gene="RIN2"
                     /locus_tag="RP5-1175I6.1-003"
                     /standard_name="OTTHUMP00000030372"
                     /note="match: proteins: Sw:Q8WYP3 Sw:Q9D684"
                     /codon_start=1
                     /product="Ras and Rab interactor 2"
                     /protein_id="CAI23572.1"
                     /db_xref="GI:56204801"
                     /db_xref="GOA:Q5TFT8"
                     /db_xref="InterPro:IPR000159"
                     /db_xref="InterPro:IPR000980"
                     /db_xref="InterPro:IPR003123"
                     /db_xref="UniProt/TrEMBL:Q5TFT8"
                     /translation="MLDSFSQESTLPFREARKRTSFQPVQVWRNFTASQTTESPACSG
                     ASLGEMTAWTMGARGLDKRGSFFKLIDTIASEIGELKQEMVRTDVNLENGLEPAETHS
                     MVRHKDGGYSEEEDVKTCARDSGYDSLSNRLSILDRLLHTHPIWLQLSLSEEEAAEVL
                     QAQPPGIFLVHKSTKMQKKVLSLRLPCEFGAPLKEFAIKESTYTFSLEGSGISFADLF
                     RLIAFYCISRDVLPFTLKLPYAISTAKSEAQLEELAQMGLNFWSSPADSKPPNLPPPH
                     RPLSSDGVCPASLRQLCLINGVHSIKTRTPSELECSQTNGALCFINPLFLKVHSQDLS
                     GGLKRPSTRTPNANGTERTRSPPPRPPPPAINSLHTSPRLARTETQTSMPETVNHNKH
                     GNVALPGTKPTPIPPPRLKKQASFLEAEGGAKTLSGGRPGAGPELELGTAGSPGGAPP
                     EAAPGDCTRAPPPSSESRPPCHGGRQRLSDMSISTSSSDSLEFDRSMPLFGYEADTNS
                     SLEDYEGESDQETMAPPIKSKKKRSSSFVLPKLVKSQLQKVSGVFSSFMTPEKRMVRR
                     IAELSRDKCTYFGCLVQDYVSFLQENKECHVSSTDMLQTIRQFMTQVKNYLSQSSELD
                     PPIESLIPEDQIDVVLEKAMHKCILKPLKGHVEAMLKDFHMADGSWKQLKENLQLVRQ
                     RNPQELGVFAPTPDFVDVEKIKVKFMTMQKMYSPEKKVMLLLRVCKLIYTVMENNSGR
                     MYGADDFLPVLTYVIAQCDMLELDTEIEYMMELLDPSLLHGEGGYYLTSAYGALSLIK
                     NFQEEQAARLLSSETRDTLRQWHKRRTTNRTIPSVDDFQNYLRVAFQEVNSGCTGKTL
                     LVRPYITTEDVCQICAEKFKVGDPEEYSLFLFVDETWQQLAEDTYPQKIKAELHSRPQ
                     PHIFHFVYKRIKNDPYGIIFQNGEEDLTTS"
     gene            join(66854. .66903,67045. .67137,AL049538.9:12596. .12696,
                     AL049538.9:34112. .34304)
                     /gene="RIN2"
                     /locus_tag="RP5-1175I6.1-002"
     mRNA            join(66854. .66903,67045. .67137,AL049538.9:12596. .12696,
                     AL049538.9:34112. .34304)
                     /gene="RIN2"
                     /locus_tag="RP5-1175I6.1-002"
                     /product="Ras and Rab interactor 2"
                     /note="match: ESTs: Em:BB854465.1 Em:BB860558.1"
     gene            join(66866. .66903,67045. .67137,AL049538.9:12596. .12696,
                     AL049538.9:12936. .13104,AL049538.9:34112. .34304,
                     AL049538.9:38197. .38308,AL049538.9:42449. .42521,
                     AL049538.9:48335. .48426)
                     /gene="RIN2"
                     /locus_tag="RP5-1175I6.1-005"
     mRNA            join(66866. .66903,67045. .67137,AL049538.9:12596. .12696,
                     AL049538.9:12936. .13104,AL049538.9:34112. .34304,
                     AL049538.9:38197. .38308,AL049538.9:42449. .42521,
                     AL049538.9:48335. .48426)
                     /gene="RIN2"
                     /locus_tag="RP5-1175I6.1-005"
                     /product="Ras and Rab interactor 2"
                     /note="match: ESTs: Em:AL709793.1 Em:AW910942.1
                     Em:BY160379.1 Em:BY169905.1 Em:BY185784.1 Em:BY743618.1
                     Em:BY745234.1 Em:BY747585.1"
     gene            join(66889. .67137,AL049538.9:12596. .12696,
                     AL049538.9:12936. .13047)
                     /gene="RIN2"
                     /locus_tag="RP5-1175I6.1-006"
     mRNA            join(66889. .67137,AL049538.9:12596. .12696,
```

```
                     AL049538.9:12936. .13047)
                     /gene="RIN2"
                     /locus_tag="RP5-1175I6.1-006"
                     /product="Ras and Rab interactor 2"
                     /note="match: ESTs: Em:BB569261.1 Em:BB849202.1"
     gene            join(67045. .67137,73239. .73426)
                     /gene="RIN2"
                     /locus_tag="RP5-1175I6.1-004"
     mRNA            join(67045. .67137,73239. .73426)
                     /gene="RIN2"
                     /locus_tag="RP5-1175I6.1-004"
                     /product="Ras and Rab interactor 2"
                     /note="match: ESTs: Em:AA953394.1 Em:AW272393.1"
     misc_feature    97273. .97318
                     /note="Single clone region"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 8;  Length 100080;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db      72005 TCAGGCATGAGCCAGCA 72021


RESULT 26
AL603749
LOCUS       AL603749              102313 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-133N1 on chromosome 1 Contains
            two novel proteins similar to preferentially expressed antigen in
            melanoma (PRAME) and the 3' end of a novel leucine rich repeat
            domain containing protein, complete sequence.
ACCESSION   AL603749
VERSION     AL603749.6  GI:16973163
KEYWORDS    HTG; leucine rich repeat; PRAME.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 102313)
  AUTHORS   Whitehead,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 16, 2001 this sequence version replaced gi:15723828.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-133N1 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES    Location/Qualifiers
     source          1. .102313
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-133N1"
                     /clone_lib="RPCI-11.1"
     misc_feature    1
                     /note="Clone_left_end: RP11-133N1"
     gene            6447. .9423
                     /locus_tag="RP11-133N1.1-001"
     mRNA            join(6447. .6763,7160. .7738,8763. .9423)
                     /locus_tag="RP11-133N1.1-001"
                     /product="novel protein similar to preferentially
                     expressed antigen in melanoma (PRAME)"
     CDS             join(6477. .6763,7160. .7738,8763. .9321)
                     /locus_tag="RP11-133N1.1-001"
                     /standard_name="OTTHUMP00000009637"
                     /codon_start=1
                     /product="novel protein similar to preferentially
                     expressed antigen in melanoma (PRAME)"
                     /protein_id="CAH70435.1"
                     /db_xref="GI:55665601"
                     /db_xref="UniProt/TrEMBL:Q5VTA0"
                     /translation="MSLQSPSRLLELAGQSLLRNQFLTIFILDELPREVFPLMFMEAS
                     SMRHFEALKLMVQAWPFLRLPLGSLMKTPHLETLQAVLKGLDTLLAQKLRPRRWKLQV
                     LDLRDVDGNFWTIWSGARALSCSPEAMSKRQTVEDYPRTGEHQPLKVFIDLCQKESTL
                     DECLSYLCRWIHYRRGLVHLCCNKVQNYSMPTSSFRNLLKRVYPDSIQELEIKRKCSL
                     NKTGKFAPYLSQMSNLRKLFLAFGYDDELYVSGQQQFVPDLDCPFLCLYYPQMLYIRK
                     ISNIKEHLEHLLRCLKNPLGTFIFCHAYLADQDMECLSQYPSLSQLKELHLIHILMWT
                     TNLEPLGALLEKVAATLEILTLKDCQIQDSQLRVLLPALSRCSQLTTFYFRGNETSTN
                     ALKDLLCHTGGLSKLGLELYPAPLECLDNRGHVNWEILAPIRAELMCTLREVRQPKRI
                     FFGPIPCPSCGSWPSEKVDFHLCS"
     gene            27266. .38162
                     /locus_tag="RP11-133N1.2-001"
     mRNA            join(27266. .27360,33167. .33457,34956. .35534,37531. .38162)
                     /locus_tag="RP11-133N1.2-001"
                     /product="novel protein similar to preferentially
                     expressed antigen in melanoma (PRAME)"
                     /note="match: cDNAs: Em:AK124292.1 Em:AL049680.1
                     Em:AL049681.1 Em:AL049682.1"
     CDS             join(33171. .33457,34956. .35534,37531. .38092)
                     /locus_tag="RP11-133N1.2-001"
                     /standard_name="OTTHUMP00000009638"
                     /note="match: proteins: Sw:O60810 Sw:O60813 Sw:O95522
                     Tr:Q8IXN8 Tr:Q9UQP2"
                     /codon_start=1
                     /product="novel protein similar to preferentially
                     expressed antigen in melanoma (PRAME)"
                     /protein_id="CAH70437.1"
                     /db_xref="GI:55665603"
                     /db_xref="InterPro:IPR001611"
                     /db_xref="UniProt/TrEMBL:Q5VT98"
                     /translation="MSIRTPPRLLELAGRSLLRDEALAISTLEELPTELFPPLFMEAF
                     SRRHCEALKLMVQAWPFLRLPLGSLMKRPCPETFQAVLDGLDALLTHRVRLRRWKLQV
                     LDLQDVSENFWMVWSEAMARRCLPNAMMNRKPVQDCPRMRGQQPLTVFIDLCLKNRTL
                     DEYFTCLFLWVKQREGLVHLCCKKLKMLGMLFHNIRNILKTVNLDCIQEVEVNCNWTL
                     PVLAEFTPYLGQMRNLRKLVLSDIDSRYISPEQKKEFVTQFTTQFLKLRCLQKLYMNS
                     VSFLEGHLDQMLSCLKTSLNILAITNCVLLESDLKHLSKYPSIGQLKTLDLSGTRLAN
                     FSLVPLQVLLEKVAATLEYLDLDDCGIVDSQVNAILPALSRCFELTTFSFRGNPISTA
                     TLENLLCHTIRLNNLCLELYPAPRESYDVRGIVCRSRFAQLGAELMGRVRALREPERI
                     LFCTDYCPQCGNRSLYDLEVDRCCC"
     misc_feature    55651. .55661
                     /note="Sequence from uni-directional dGTP big dye
                     terminator reads only"
     misc_feature    85607
                     /note="Clone_right_end: RP11-584P2"
     gene            join(complement(AL354712.18:29504. .30589),
                     complement(91804. .92926))
                     /gene="RP4-597A16.1"
                     /locus_tag="RP4-597A16.1-001"
```

```
     mRNA            join(complement(AL354712.18:29504. .30589),
                     complement(91804. .92926))
                     /gene="RP4-597A16.1"
                     /locus_tag="RP4-597A16.1-001"
                     /product="novel leucine rich repeat domain containing
                     protein"
                     /note="match: ESTs: Em:AA431938.1 Em:AA719301.1
                     Em:BE618324.1 Em:BG396264.1 Em:BG472142.1 Em:BM925485.1
                     Em:BX391678.1 Em:BX396881.1 Em:BX396882.1
                     match: cDNAs: Em:AK038658.1 Em:AK045936.1 Em:AK078190.1
                     Em:BC016048.1"
     polyA_site      complement(91804)
                     /gene="RP4-597A16.1"
                     /locus_tag="RP4-597A16.1-001"
     polyA_signal    complement(91820. .91825)
                     /gene="RP4-597A16.1"
                     /locus_tag="RP4-597A16.1-001"
     polyA_signal    complement(91824. .91829)
                     /gene="RP4-597A16.1"
                     /locus_tag="RP4-597A16.1-001"
     CDS             join(complement(AL354712.18:29504. .30134),
                     complement(92673. .92926))
                     /gene="RP4-597A16.1"
                     /locus_tag="RP4-597A16.1-001"
                     /standard_name="OTTHUMP00000009643"
                     /note="match: proteins: Tr:Q8BR15 Tr:Q96B32"
                     /codon_start=1
                     /product="novel leucine rich repeat domain containing
                     protein"
                     /protein_id="CAH70436.1"
                     /db_xref="GI:55665602"
                     /db_xref="InterPro:IPR000372"
                     /db_xref="InterPro:IPR000483"
                     /db_xref="InterPro:IPR001611"
                     /db_xref="InterPro:IPR003591"
                     /db_xref="UniProt/TrEMBL:Q5VT99"
                     /translation="MRPRAPACAAAALGLCSLLLLLAPGHACPAGCACTDPHTVDCRD
                     RGLPSVPDPFPLDVRKLLVAGNRIQRIPEDFFIFYGDLVYLDFRNNSLRSLEEGTFSG
                     SAKLVFLDLSYNNLTQLGAGAFRSAGRLVKLSLANNNLVGVHEDAFETLESLQVLELN
                     DNNLRSLSVAALAALPALRSLRLDGNPWLCDCDFAHLFSWIQENASKLPKGLDEIQCS
                     LPMESRRISLRELSEASFSECRFSLSLTDLCIIIFSGVAVSIAAIISSFFLATVVQCL
                     QRCAPNKDAEDEDEDKDD"
     misc_feature    100314
                     /note="Clone_left_end: RP4-597A16"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 8;  Length 102313;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db      41207 TCAGGCATGAGCCAGCA 41223


RESULT 27
CR352288
LOCUS       CR352288              103979 bp    DNA     linear   VRT 19-MAR-2005
DEFINITION  Zebrafish DNA sequence from clone CH211-123P18 in linkage group 16,
            complete sequence.
ACCESSION   CR352288
VERSION     CR352288.9  GI:61657063
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 103979)
  AUTHORS   Gray,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-MAR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Mar 19, 2005 this sequence version replaced gi:58651979.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml
            CH211-123P18 is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1.
FEATURES             Location/Qualifiers
     source          1. .103979
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-123P18"
                     /clone_lib="CHORI-211"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 5;  Length 103979;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 TCTGGAGCTCAGGCATG 21
              |||||||||||||||||
Db      94915 TCTGGAGCTCAGGCATG 94931


RESULT 28
AC091528_1
WPCOMMENT
  Sequence split into 5 fragments  LOCUS AC091528 Accession AC091528
     Fragment Name      Begin     End
     AC091528_0             1     110000
     AC091528_1        100001     210000
     AC091528_2        200001     310000
     AC091528_3        300001     410000
     AC091528_4        400001     431617
Continuation (2 of 5) of AC091528 from base 100001 (AC091528 Homo sapiens chromosome 12

  Query Match             58.6%;  Score 17;  DB 14;  Length 110000;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
```

```
                  ||||||||||||||||||
Db      50975 TCAGGCATGAGCCAGCA 50991


RESULT 29
AC104812/c
LOCUS       AC104812               120574 bp    DNA     linear   PRI 30-APR-2005
DEFINITION  Homo sapiens BAC clone RP11-656O23 from 2, complete sequence.
ACCESSION   AC104812
VERSION     AC104812.5  GI:20128749
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 120574)
  AUTHORS   Doebber,A., Haglund,K. and Boatright,E.
  TITLE     The sequence of Homo sapiens BAC clone RP11-656O23
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 120574)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-DEC-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 120574)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 120574)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAY-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 120574)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-APR-2005) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Apr 10, 2002 this sequence version replaced gi:18464356.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0656O23
            --------------.

            NOTICE:

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
            VECTOR:  pBACe3.6

            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-419O4, 2000 bp overlap; the
            clone sequenced to the right is RP11-1152J17, 2000 bp overlap.
            Actual start of this clone is at base position 42704 of RP11-419O4;
            actual end is at base position 36755 of RP11-1152J17.

            Polymorphisms have been identified between AC068487, AC092843 and
            AC104812.  Data from AC068487 was used to finish this clone,
            AC104812.
FEATURES             Location/Qualifiers
     source          1. .120574
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /clone="RP11-656O23"
                     /clone_lib="RPCI-11"
     misc_feature    19619. .19870
                     /note="CpG_island (%GC=61.1, o/e=0.74, #CpGs=20)"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 8;  Length 120574;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TCTCTGGAGCTCAGGCA 19
              |||||||||||||||||
Db      48297 TCTCTGGAGCTCAGGCA 48281


RESULT 30
CR384099/c
LOCUS       CR384099               121857 bp    DNA     linear   VRT 09-APR-2005
DEFINITION  Zebrafish DNA sequence from clone DKEY-266I18 in linkage group 24,
            complete sequence.
ACCESSION   CR384099
VERSION     CR384099.6  GI:62460976
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 121857)
  AUTHORS   Clark,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-APR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 9, 2005 this sequence version replaced gi:61969276.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
```

```
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-266I18
            is from a Zebrafish BAC library
            VECTOR: pIndigoBAC-5.
FEATURES             Location/Qualifiers
     source          1. .121857
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-266I18"
                     /clone_lib="DanioKey"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 5;  Length 121857;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CTCTCTGGAGCTCAGGC 18
              |||||||||||||||||
Db      84712 CTCTCTGGAGCTCAGGC 84696


RESULT 31
CR339049
LOCUS       CR339049              126083 bp    DNA     linear   VRT 18-SEP-2004
DEFINITION  Zebrafish DNA sequence from clone CH211-165D12 in linkage group 3,
            complete sequence.
ACCESSION   CR339049
VERSION     CR339049.6  GI:52345338
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 126083)
  AUTHORS   Hunter,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-SEP-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Sep 18, 2004 this sequence version replaced gi:51591762.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml
            CH211-165D12 is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1.
FEATURES             Location/Qualifiers
     source          1. .126083
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-165D12"
                     /clone_lib="CHORI-211"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 5;  Length 126083;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CTCTCTGGAGCTCAGGC 18
              |||||||||||||||||
Db      94588 CTCTCTGGAGCTCAGGC 94604


RESULT 32
AL611933/c
LOCUS       AL611933              141273 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-374C13 on chromosome 1 Contains
            a eukaryotic translation elongation factor 1 delta (guanine
            nucleotide exchange protein) (EEF1D) pseudogene, complete sequence.
ACCESSION   AL611933
VERSION     AL611933.30  GI:21425229
KEYWORDS    HTG; EEF1D; translation elongation factor.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 141273)
  AUTHORS   Peck,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 13, 2002 this sequence version replaced gi:21261816.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-374C13 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
```

```
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .141273
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-374C13"
                     /clone_lib="RPCI-11.2"
     misc_feature    1
                     /note="Clone_left_end: RP11-374C13"
     misc_feature    15302. .15460
                     /note="Single read sequenced with dGTP big dye terminator
                     chemistry. Assembly confirmed by restriction digest data."
     misc_feature    15461
                     /note="Tandem repeat. Forced join. Gap size estimated to
                     be approximately 500bp by restriction digest data."
     misc_feature    23111. .23359
                     /note="Assembly confirmed by restriction digest data."
     misc_feature    23113. .23116
                     /note="Weak data."
     misc_feature    23360
                     /note="Tandem repeat. Forced join. Gap size estimated to
                     be approximately 1000bp by restriction digest data."
     misc_feature    41419. .41420
                     /note="Weak data."
     misc_feature    41437. .41449
                     /note="Single clone region. Assembly confirmed by
                     restriction digest data."
     misc_feature    41450
                     /note="Tandem repeat. Forced join. Gap size estimated to
                     be approximately 400bp by restriction digest data."
     gene            complement(54257. .54628)
                     /locus_tag="RP11-374C13.1-001"
                     /pseudo
     CDS             complement(54257. .54628)
                     /locus_tag="RP11-374C13.1-001"
                     /note="match: proteins: Sw:P29692 Sw:P57776 Tr:Q80T06
                     Tr:Q91VK2 Tr:Q96I38 Tr:Q9BW34"
                     /pseudo
                     /codon_start=1
                     /product="eukaryotic translation elongation factor 1 delta
                     (guanine nucleotide exchange protein) (EEF1D) pseudogene"
     misc_feature    65497. .65641
                     /note="Sequence from clone PCR only. Assembly confirmed by
                     restriction digest data."
     misc_feature    139274
                     /note="Clone_left_end: RP11-168B8"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 8;  Length 141273;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db      93586 TCAGGCATGAGCCAGCA 93570


RESULT 33
AL591936/c
LOCUS       AL591936              147431 bp    DNA     linear   ROD 01-JUN-2003
DEFINITION  Mouse DNA sequence from clone RP23-28B10 on chromosome 2, complete
            sequence.
ACCESSION   AL591936
VERSION     AL591936.12  GI:31335545
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 147431)
  AUTHORS   Sycamore,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 2, 2003 this sequence version replaced gi:21211801.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep --------------
            Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            RP23-28B10 is from the RPCI-23 Mouse BAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            Sequence from the Mouse Genome Sequencing Consortium whole genome
            shotgun may have been used to confirm this sequence.  Sequence data
            from the whole genome shotgun alone has only been used where it has
            a phred quality of at least 30.
FEATURES             Location/Qualifiers
     source          1. .147431
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="2"
                     /clone="RP23-28B10"
                     /clone_lib="RPCI-23"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 9;  Length 147431;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 TGGAGCTCAGGCATGAG 23
              |||||||||||||||||
Db      83748 TGGAGCTCAGGCATGAG 83732


RESULT 34
AC069440/c
LOCUS       AC069440              149817 bp    DNA     linear   HTG 10-NOV-2000
DEFINITION  Homo sapiens chromosome 3 clone RP11-1022P15, *** SEQUENCING IN
```

```
            PROGRESS ***, 52 unordered pieces.
ACCESSION   AC069440
VERSION     AC069440.7  GI:11128260
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 149817)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
            Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
            Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
            Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
            Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Morgan,M., Morris,S., Moser,M., Neal,D., Newtson,J., Newtson,N.,
            Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
            Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
            Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
            Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
            Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
            Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
            Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N.,
            Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
            Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
            Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
            Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.
            and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 149817)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAY-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 9, 2000 this sequence version replaced gi:10180136.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: HMLE
                Center clone name: RP11-1022P15
            -------------- Summary Statistics
                Sequencing vector: M13; L08821
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 84144 bases at least Q40
                Consensus quality: 109304 bases at least Q30
                Consensus quality: 126815 bases at least Q20
                Estimated insert size: 127243; sum-of-contigs estimation
                Quality coverage: 0x in Q20 bases; agarose-fp estimation
                Quality coverage: 2x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 52 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     6430: contig of 6430 bp in length
            *    6431     6530: gap of unknown length
            *    6531    12107: contig of 5577 bp in length
            *   12108    12207: gap of unknown length
            *   12208    17155: contig of 4948 bp in length
            *   17156    17255: gap of unknown length
            *   17256    23677: contig of 6422 bp in length
            *   23678    23777: gap of unknown length
            *   23778    29891: contig of 6114 bp in length
            *   29892    29991: gap of unknown length
            *   29992    34509: contig of 4518 bp in length
            *   34510    34609: gap of unknown length
            *   34610    39930: contig of 5321 bp in length
            *   39931    40030: gap of unknown length
            *   40031    45359: contig of 5329 bp in length
            *   45360    45459: gap of unknown length
            *   45460    51223: contig of 5764 bp in length
            *   51224    51323: gap of unknown length
            *   51324    55094: contig of 3771 bp in length
            *   55095    55194: gap of unknown length
            *   55195    60265: contig of 5071 bp in length
            *   60266    60365: gap of unknown length
            *   60366    65357: contig of 4992 bp in length
            *   65358    65457: gap of unknown length
            *   65458    68265: contig of 2808 bp in length
            *   68266    68365: gap of unknown length
            *   68366    71160: contig of 2795 bp in length
            *   71161    71260: gap of unknown length
            *   71261    74282: contig of 3022 bp in length
            *   74283    74382: gap of unknown length
            *   74383    76728: contig of 2346 bp in length
            *   76729    76828: gap of unknown length
            *   76829    80232: contig of 3404 bp in length
            *   80233    80332: gap of unknown length
            *   80333    83471: contig of 3139 bp in length
            *   83472    83571: gap of unknown length
            *   83572    86081: contig of 2510 bp in length
            *   86082    86181: gap of unknown length
            *   86182    88957: contig of 2776 bp in length
            *   88958    89057: gap of unknown length
            *   89058    91534: contig of 2477 bp in length
            *   91535    91634: gap of unknown length
            *   91635    94187: contig of 2553 bp in length
            *   94188    94287: gap of unknown length
            *   94288    96737: contig of 2450 bp in length
            *   96738    96837: gap of unknown length
            *   96838    99244: contig of 2407 bp in length
            *   99245    99344: gap of unknown length
            *   99345   101231: contig of 1887 bp in length
            *  101232   101331: gap of unknown length
            *  101332   104255: contig of 2924 bp in length
            *  104256   104355: gap of unknown length
            *  104356   105819: contig of 1464 bp in length
            *  105820   105919: gap of unknown length
            *  105920   107463: contig of 1544 bp in length
            *  107464   107563: gap of unknown length
            *  107564   109580: contig of 2017 bp in length
            *  109581   109680: gap of unknown length
            *  109681   111751: contig of 2071 bp in length
```

```
            *   111752   111851: gap of unknown length
            *   111852   113921: contig of 2070 bp in length
            *   113922   114021: gap of unknown length
            *   114022   116526: contig of 2505 bp in length
            *   116527   116626: gap of unknown length
            *   116627   118510: contig of 1884 bp in length
            *   118511   118610: gap of unknown length
            *   118611   120242: contig of 1632 bp in length
            *   120243   120342: gap of unknown length
            *   120343   122674: contig of 2332 bp in length
            *   122675   122774: gap of unknown length
            *   122775   124266: contig of 1492 bp in length
            *   124267   124366: gap of unknown length
            *   124367   125990: contig of 1624 bp in length
            *   125991   126090: gap of unknown length
            *   126091   127132: contig of 1042 bp in length
            *   127133   127232: gap of unknown length
            *   127233   129519: contig of 2287 bp in length
            *   129520   129619: gap of unknown length
            *   129620   132039: contig of 2420 bp in length
            *   132040   132139: gap of unknown length
            *   132140   133632: contig of 1493 bp in length
            *   133633   133732: gap of unknown length
            *   133733   136088: contig of 2356 bp in length
            *   136089   136188: gap of unknown length
            *   136189   137639: contig of 1451 bp in length
            *   137640   137739: gap of unknown length
            *   137740   138800: contig of 1061 bp in length
            *   138801   138900: gap of unknown length
            *   138901   140325: contig of 1425 bp in length
            *   140326   140425: gap of unknown length
            *   140426   141944: contig of 1519 bp in length
            *   141945   142044: gap of unknown length
            *   142045   143164: contig of 1120 bp in length
            *   143165   143264: gap of unknown length
            *   143265   144684: contig of 1420 bp in length
            *   144685   144784: gap of unknown length
            *   144785   146066: contig of 1282 bp in length
            *   146067   146166: gap of unknown length
            *   146167   147480: contig of 1314 bp in length
            *   147481   147580: gap of unknown length
            *   147581   148690: contig of 1110 bp in length
            *   148691   148790: gap of unknown length
            *   148791   149817: contig of 1027 bp in length.
FEATURES             Location/Qualifiers
     source          1. .149817
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /clone="RP11-1022P15"
     gap             6431. .6530

  Query Match             58.6%;  Score 17;  DB 14;  Length 149817;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          13 TCAGGCATGAGCCAGCA 29
               |||||||||||||||||
Db       44738 TCAGGCATGAGCCAGCA 44722


RESULT 35
AF186190/c
LOCUS       AF186190               152405 bp    DNA     linear   PRI 10-JUL-2002
DEFINITION  Homo sapiens chromosome 8 clone CTC-369M3 map 8q24.3 , complete
            sequence.
ACCESSION   AF186190
VERSION     AF186190.4  GI:21724079
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 152405)
  AUTHORS   Blechschmidt,K., Schattevoy,R., Baumgart,C. and Rosenthal,A.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 152405)
  AUTHORS   Blechschmidt,K., Kalaydjieva,L., Goodman,R., Gresham,D., Baas,F.,
            Jonge,Rd., Schilhabel,M., Schattevoy,R., Baumgart,C., Menzel,U. and
            Rosenthal,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-SEP-1999) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
REFERENCE   3  (bases 1 to 152405)
  AUTHORS   Genome Sequencing Center Jena.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAY-2000) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenberstr. 11, Jena 07745, Germany
REFERENCE   4  (bases 1 to 152405)
  AUTHORS   Genome Sequencing Center Jena.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-2000) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenberstr. 11, Jena 07745, Germany
REFERENCE   5  (bases 1 to 152405)
  AUTHORS   Lagemann,D. and Platzer,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2002) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
COMMENT     On Jul 10, 2002 this sequence version replaced gi:11095448.
            -------------- Genome Center
                 Center: Insitute of Molecular Biotechnology
                 Center code: IMB
                 Web site: http://genome.imb-jena.de/
                 Contact: gscj-submit@genome.imb-jena.de
            -------------- Project Information
                 Center project name: H264
                 Center clone name: CTC-369M3
            -------------- Summary Statistics
                 Sequencing vector: pUC18; 100% of reads
                 Chemistry: Dye-terminator Big Dye; 100% of reads
                 Assembly program: Phrap; version 0.990329
                 Consensus quality: 151415 bases at least Q40
                 Consensus quality: 152179 bases at least Q30
                 Consensus quality: 152325 bases at least Q20
                 Quality coverage: 11.45x
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one pUC18 subclone; and the
            assembly was confirmed by restriction digest.
            --------------
            Sequence Quality Assessment:
             This entry has been annotated with sequence quality
             estimates computed by the Phrap assembly program.
             All manually edited bases have been reduced to quality zero.
             Quality levels above 40 are expected to have less than
             1 error in 10,000 bp.
             Base-by-base quality values are not generally visible from the
             GenBank flat file format but are available as part
             of this entry's ASN.1 file.
            ----------.
FEATURES             Location/Qualifiers
     source          1. .152405
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8q24.3 , complete sequence."
                     /clone="CTC-369M3"
     misc_feature    5649
                     /note="low quality region , CTC-369M3"
```

```
misc_feature    28279
                /note="low quality region , CTC-369M3"
misc_feature    28432
                /note="low quality region , CTC-369M3"
misc_feature    28445. .28815
                /note="single stranded/single chemistry region"
misc_feature    28640
                /note="low quality region , CTC-369M3"
misc_feature    28703
                /note="low quality region , CTC-369M3"
misc_feature    42911. .43061
                /note="single stranded/single chemistry region"
misc_feature    42998
                /note="low quality region , CTC-369M3"
misc_feature    43045
                /note="low quality region , CTC-369M3"
misc_feature    43051
                /note="low quality region , CTC-458A3"
misc_feature    43380
                /note="low quality region , CTC-369M3"
variation       43765
                /note="deleted in clone: CTC-458A3"
                /replace=""
misc_feature    51031
                /note="low quality region , CTC-458A3"
misc_feature    51219. .51767
                /note="single stranded/single chemistry region"
misc_feature    51334
                /note="low quality region , CTC-458A3"
misc_feature    51496
                /note="low quality region , CTC-369M3"
variation       54126. .54129
                /note="deleted in clone: CTC-369M3"
                /replace=""
misc_feature    59349. .59437
                /note="single stranded/single chemistry region"
misc_feature    61624. .61861
                /note="single stranded/single chemistry region"
misc_feature    61849
                /note="low quality region , CTC-458A3"
unsure          61849
                /note="CTC-458A3"
misc_feature    61862. .62251
                /note="single stranded/single chemistry region"
misc_feature    61972. .62571
                /note="pcr product sequence only , CTC-458A3"
misc_feature    62252. .62415
                /note="single stranded/single chemistry region"
variation       62257. .62258
                /note="AA substituted in clone: CTC-369M3"
                /replace="tg"
variation       62264
                /note="A substituted in clone: CTC-458A3"
                /replace="t"
variation       62280
                /note="T substituted in clone: CTC-458A3"
                /replace="c"
variation       62281
                /note="G substituted in clone: CTC-458A3"
                /replace="a"
misc_feature    62381
                /note="low quality region , CTC-458A3"
misc_feature    62412. .62415
                /note="low quality region , CTC-458A3"
misc_feature    62572. .62661
                /note="pcr product sequence only , CTC-369M3"
misc_feature    63555. .63719
                /note="single stranded/single chemistry region"
misc_feature    70669. .70931
                /note="single stranded/single chemistry region"
misc_feature    70932. .71302
                /note="single stranded/single chemistry region"
misc_feature    71231
                /note="low quality region , CTC-369M3"
misc_feature    71303. .71584
                /note="single stranded/single chemistry region"
misc_feature    71333
                /note="low quality region , CTC-369M3"
misc_feature    74692. .75005
                /note="single stranded/single chemistry region"
misc_feature    75006. .75135
                /note="single stranded/single chemistry region"
misc_feature    75044
                /note="low quality region , CTC-458A3"
misc_feature    75752. .75829
                /note="single stranded/single chemistry region"
misc_feature    75830. .75913
                /note="single stranded/single chemistry region"
misc_feature    77345
                /note="low quality region , CTC-458A3"
misc_feature    81361. .81655
                /note="single stranded/single chemistry region"
misc_feature    89327
                /note="low quality region , CTC-369M3"
misc_feature    89336
                /note="low quality region , CTC-369M3"
misc_feature    89338
                /note="low quality region , CTC-369M3"
unsure          89338
                /note="CTC-369M3"
misc_feature    89370
                /note="low quality region , CTC-369M3"
unsure          89370
                /note="CTC-369M3"
misc_feature    89507
                /note="low quality region , CTC-369M3"
misc_feature    90758
                /note="low quality region , CTC-369M3"
misc_feature    90790
                /note="low quality region , CTC-369M3"
misc_feature    98691
                /note="low quality region , CTC-369M3"
misc_feature    98694
                /note="low quality region , CTC-369M3"
misc_feature    98732
                /note="low quality region , CTC-369M3"
misc_feature    98734
                /note="low quality region , CTC-369M3"

  Query Match             58.6%;  Score 17;  DB 8;  Length 152405;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db      98806 TCAGGCATGAGCCAGCA 98790


RESULT 36
AC138625
LOCUS       AC138625              152884 bp    DNA     linear   PRI 14-JAN-2003
DEFINITION  Homo sapiens chromosome 16 clone RP11-23E19, complete sequence.
ACCESSION   AC138625
VERSION     AC138625.1  GI:27733930
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 152884)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 152884)
```

```
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-2003) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center and Los Alamos
            National Laboratory
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.
FEATURES             Location/Qualifiers
     source          1. .152884
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="RP11-23E19"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 8;  Length 152884;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAGC 24
              |||||||||||||||||
Db      48052 GGAGCTCAGGCATGAGC 48068


RESULT 37
CR854849/c
LOCUS       CR854849              153012 bp    DNA     linear   PRI 22-MAY-2005
DEFINITION  Human DNA sequence from clone RP13-79M23 on chromosome 1, complete
            sequence.
ACCESSION   CR854849
VERSION     CR854849.7  GI:66392951
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 153012)
  AUTHORS   Barlow,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On May 22, 2005 this sequence version replaced gi:57863715.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP13-79M23 is from the library RPCI-13.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .153012
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP13-79M23"
                     /clone_lib="RPCI-13.1"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 8;  Length 153012;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db       2319 TCAGGCATGAGCCAGCA 2303


RESULT 38
AC112229
LOCUS       AC112229              154279 bp    DNA     linear   PRI 08-APR-2005
DEFINITION  Homo sapiens BAC clone RP13-1039J1 from 2, complete sequence.
ACCESSION   AC112229
VERSION     AC112229.4  GI:22758607
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 154279)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   2  (bases 1 to 154279)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUL-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 154279)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 154279)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-SEP-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 154279)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-APR-2005) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Sep 7, 2002 this sequence version replaced gi:21747810.
FEATURES             Location/Qualifiers
     source          1. .154279
```

```
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2"
                     /clone="RP13-1039J1"
                     /clone_lib="RPCI-13"
     misc_feature    36786. .37693
                     /note="CpG island (%GC=71.8, o/e=0.74, #CpGs=96)"
     misc_feature    76110. .77781
                     /note="CpG island (%GC=65.9, o/e=0.67, #CpGs=139)"
     misc_feature    153268. .154028
                     /note="CpG island (%GC=66.2, o/e=0.72, #CpGs=53)"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 8;  Length 154279;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 GAGCTCAGGCATGAGCC 25
              |||||||||||||||||
Db     114426 GAGCTCAGGCATGAGCC 114442


RESULT 39
AC138879/c
LOCUS       AC138879              154729 bp    DNA     linear   HTG 21-JAN-2003
DEFINITION  Homo sapiens chromosome 16 clone RP11-268J7, WORKING DRAFT
            SEQUENCE.
ACCESSION   AC138879
VERSION     AC138879.1  GI:27805291
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 154729)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 16
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 154729)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JAN-2003) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 509500
            Center clone name: RPCI-11_268J7
            -------
            Summary Statistics
            Consensus quality: 154729 bases at least Q40
            Consensus quality: 154729 bases at least Q30
            Consensus quality: 154729 bases at least Q20
            Estimated insert size: 180000; agarose-fp estimation
            Estimated insert size: 154729; sum-of-contigs estimation
            Quality coverage:  12.71 in Q20 bases; agarose-fp estimation
            Quality coverage:  14.78 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *         1   154729: contig of 154729 bp in length.
FEATURES            Location/Qualifiers
     source         1. .154729
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /chromosome="16"
                    /clone="RP11-268J7"
                    /clone_lib="RPCI human BAC library 11"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 14;  Length 154729;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAGC 24
              |||||||||||||||||
Db      39571 GGAGCTCAGGCATGAGC 39555


RESULT 40
AC023236
LOCUS       AC023236              154732 bp    DNA     linear   HTG 06-SEP-2000
DEFINITION  Homo sapiens chromosome 3 clone RP11-382A21, WORKING DRAFT
            SEQUENCE, 36 unordered pieces.
ACCESSION   AC023236
VERSION     AC023236.11  GI:9438287
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 154732)
  AUTHORS   Muzny,D.M., Adams,C., Bailey,M., Barbaria,J., Blankenburg,K.,
            Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
            Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
            Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
            Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
            Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
            Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
            Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
            Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
            Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
            Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R.,
            Tabor,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wahbah,M.,
            Watlington,S., Weinstock,G., Weinstock,I.R., Williamson,A.,
            Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D. and
            Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 154732)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-FEB-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jul 25, 2000 this sequence version replaced gi:8571496.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: HAER
                Center clone name: RP11-382A21
            -------------- Summary Statistics
                Sequencing vector: M13; L08821
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 110591 bases at least Q40
```

```
    Consensus quality: 131981 bases at least Q30
    Consensus quality: 140244 bases at least Q20
    Estimated insert size: 140907; sum-of-contigs estimation
    Quality coverage: 0x in Q20 bases; agarose-fp estimation
    Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
--------------
* NOTE: Estimated insert size may differ from sequence length
*   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 36 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1    16063: contig of 16063 bp in length
*    16064    16163: gap of unknown length
*    16164    33051: contig of 16888 bp in length
*    33052    33151: gap of unknown length
*    33152    42833: contig of 9682 bp in length
*    42834    42933: gap of unknown length
*    42934    51643: contig of 8710 bp in length
*    51644    51743: gap of unknown length
*    51744    58186: contig of 6443 bp in length
*    58187    58286: gap of unknown length
*    58287    64740: contig of 6454 bp in length
*    64741    64840: gap of unknown length
*    64841    72968: contig of 8128 bp in length
*    72969    73068: gap of unknown length
*    73069    80043: contig of 6975 bp in length
*    80044    80143: gap of unknown length
*    80144    85494: contig of 5351 bp in length
*    85495    85594: gap of unknown length
*    85595    91128: contig of 5534 bp in length
*    91129    91228: gap of unknown length
*    91229    96066: contig of 4838 bp in length
*    96067    96166: gap of unknown length
*    96167   101533: contig of 5367 bp in length
*   101534   101633: gap of unknown length
*   101634   106721: contig of 5088 bp in length
*   106722   106821: gap of unknown length
*   106822   109023: contig of 2202 bp in length
*   109024   109123: gap of unknown length
*   109124   111925: contig of 2802 bp in length
*   111926   112025: gap of unknown length
*   112026   114559: contig of 2534 bp in length
*   114560   114659: gap of unknown length
*   114660   117368: contig of 2709 bp in length
*   117369   117468: gap of unknown length
*   117469   120143: contig of 2675 bp in length
*   120144   120243: gap of unknown length
*   120244   123405: contig of 3162 bp in length
*   123406   123505: gap of unknown length
*   123506   127090: contig of 3585 bp in length
*   127091   127190: gap of unknown length
*   127191   129044: contig of 1854 bp in length
*   129045   129144: gap of unknown length
*   129145   131922: contig of 2778 bp in length
*   131923   132022: gap of unknown length
*   132023   134234: contig of 2212 bp in length
*   134235   134334: gap of unknown length
*   134335   135891: contig of 1557 bp in length
*   135892   135991: gap of unknown length
*   135992   137278: contig of 1287 bp in length
*   137279   137378: gap of unknown length
*   137379   139397: contig of 2019 bp in length
*   139398   139497: gap of unknown length
*   139498   140592: contig of 1095 bp in length
*   140593   140692: gap of unknown length
*   140693   142276: contig of 1584 bp in length
*   142277   142376: gap of unknown length
*   142377   143937: contig of 1561 bp in length
*   143938   144037: gap of unknown length
*   144038   145268: contig of 1231 bp in length
*   145269   145368: gap of unknown length
*   145369   147144: contig of 1776 bp in length
*   147145   147244: gap of unknown length
*   147245   148728: contig of 1484 bp in length
*   148729   148828: gap of unknown length
*   148829   150426: contig of 1598 bp in length
*   150427   150526: gap of unknown length
*   150527   152015: contig of 1489 bp in length
*   152016   152115: gap of unknown length
*   152116   153552: contig of 1437 bp in length
*   153553   153652: gap of unknown length
*   153653   154732: contig of 1080 bp in length.
FEATURES             Location/Qualifiers
     source          1. .154732
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /clone="RP11-382A21"
     gap             16064. .16163
                     /estimated_length=unknown
     gap             33052. .33151
                     /estimated_length=unknown
     gap             42834. .42933
                     /estimated_length=unknown
     gap             51644. .51743
                     /estimated_length=unknown
     gap             58187. .58286
                     /estimated_length=unknown
     gap             64741. .64840
                     /estimated_length=unknown
     gap             72969. .73068
                     /estimated_length=unknown
     gap             80044. .80143
                     /estimated_length=unknown
     gap             85495. .85594
                     /estimated_length=unknown
     gap             91129. .91228
                     /estimated_length=unknown
     gap             96067. .96166
                     /estimated_length=unknown
     gap             101534. .101633
                     /estimated_length=unknown
     gap             106722. .106821
                     /estimated_length=unknown
     gap             109024. .109123
                     /estimated_length=unknown
     gap             111926. .112025
                     /estimated_length=unknown
     gap             114560. .114659
                     /estimated_length=unknown
     gap             117369. .117468
                     /estimated_length=unknown
     gap             120144. .120243
                     /estimated_length=unknown
     gap             123406. .123505
                     /estimated_length=unknown
     gap             127091. .127190
                     /estimated_length=unknown
     gap             129045. .129144
                     /estimated_length=unknown
     gap             131923. .132022
                     /estimated_length=unknown
     gap             134235. .134334
                     /estimated_length=unknown
     gap             135892. .135991
                     /estimated_length=unknown
     gap             137279. .137378
                     /estimated_length=unknown
     gap             139398. .139497
                     /estimated_length=unknown
```

```
    gap             140593. .140692

  Query Match             58.6%;  Score 17;  DB 14;  Length 154732;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          13 TCAGGCATGAGCCAGCA 29
               |||||||||||||||||
Db       59602 TCAGGCATGAGCCAGCA 59618


RESULT 41
AC141284/c
LOCUS       AC141284               156826 bp    DNA     linear   HTG 11-MAR-2003
DEFINITION  Homo sapiens chromosome 16 clone RP11-349F11, WORKING DRAFT
            SEQUENCE, 7 unordered pieces.
ACCESSION   AC141284
VERSION     AC141284.1  GI:28913064
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 156826)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 16
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 156826)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-2003) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 540512
            Center clone name: RPCI-11_349F11
            -------
            Summary Statistics
            Consensus quality: 154915 bases at least Q40
            Consensus quality: 155492 bases at least Q30
            Consensus quality: 155780 bases at least Q20
            Estimated insert size: 175000; agarose-fp estimation
            Estimated insert size: 156226; sum-of-contigs estimation
            Quality coverage:  5.68 in Q20 bases; agarose-fp estimation
            Quality coverage:  6.36 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 7 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     4363: contig of 4363 bp in length
            *     4364     4463: gap of unknown length
            *     4464    10572: contig of 6109 bp in length
            *    10573    10672: gap of unknown length
            *    10673    16705: contig of 6033 bp in length
            *    16706    16805: gap of unknown length
            *    16806    23420: contig of 6615 bp in length
            *    23421    23520: gap of unknown length
            *    23521    62938: contig of 39418 bp in length
            *    62939    63038: gap of unknown length
            *    63039   109057: contig of 46019 bp in length
            *   109058   109157: gap of unknown length
            *   109158   156826: contig of 47669 bp in length.
FEATURES             Location/Qualifiers
     source          1. .156826
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="RP11-349F11"
                     /clone_lib="RPCI human BAC library 11"
     gap             4364. .4463
                     /estimated_length=unknown
     gap             10573. .10672
                     /estimated_length=unknown
     gap             16706. .16805
                     /estimated_length=unknown
     gap             23421. .23520
                     /estimated_length=unknown
     gap             62939. .63038
                     /estimated_length=unknown
     gap             109058. .109157
                     /estimated_length=unknown
ORIGIN

  Query Match             58.6%;  Score 17;  DB 14;  Length 156826;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           9 GAGCTCAGGCATGAGCC 25
               |||||||||||||||||
Db       88379 GAGCTCAGGCATGAGCC 88363


RESULT 42
AC069421/c
LOCUS       AC069421               157289 bp    DNA     linear   PRI 12-JUN-2002
DEFINITION  Homo sapiens 3 BAC RP11-528A4 (Roswell Park Cancer Institute Human
            BAC Library) complete sequence.
ACCESSION   AC069421
VERSION     AC069421.9  GI:19033389
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 157289)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
            Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N.,
            Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N.,
            Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
```

```
            Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
            Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
            Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E.,
            Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
            Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
            Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
            Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
            Warren,R., Washington,C., Watlington,S., Williams,G.,
            Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
            Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstock,G. and
            Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 157289)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAY-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 157289)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 157289)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Mar 1, 2002 this sequence version replaced gi:18481945.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

            CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
               STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
               Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
               Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
            continuity across the splice junctions. Sequences that are not
            identical matches are annotated as similar.

            SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
            standard of double strand coverage with a minimum of 2 clones and 2
            reads with no ambiguities or 2 chemistries with a minimum of 2
            clones and 3 reads with no ambiguities. If the sequence quality for
            a region does not meet this standard, it will be indicated in the
            annotation as Low Coverage.

            QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
            standards - estimated error rate less than 1 per 10,000 bases.
            Reports of lowest quality individual bases and measures of base
            quality are listed below. Description of the metrics can be found
            at URL:
            http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

            QUALSTAT-REPORT.
FEATURES            Location/Qualifiers
     source         1. .157289
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /chromosome="3"
                    /clone="RP11-528A4"
     STS            353. .632
                    /standard_name="81567"
     STS            546. .672
                    /standard_name="21796"
     STS            546. .672
                    /standard_name="31586"
     STS            1391. .1553
                    /standard_name="70071"
     STS            1639. .1736
                    /standard_name="11888"
     repeat_region  complement(2828. .3117)
                    /rpt_family="AluY"
     repeat_region  complement(4603. .4715)
                    /rpt_family="MIR"
     repeat_region  complement(5168. .5466)
                    /rpt_family="AluJo"
     STS            6645. .6848
                    /standard_name="64408"
     repeat_region  7095. .7249
                    /rpt_family="MER102"
     repeat_region  complement(7409. .7595)
                    /rpt_family="MIR"
     repeat_region  complement(7835. .8198)
                    /rpt_family="MER47A"
     repeat_region  8522. .8930
                    /rpt_family="L2"
     repeat_region  complement(8968. .9282)
                    /rpt_family="AluSx"
     repeat_region  9864. .10096
                    /rpt_family="Charlie1a"
     repeat_region  10128. .11036
                    /rpt_family="Charlie1a"
     repeat_region  complement(11037. .11353)
                    /rpt_family="AluSq"
     repeat_region  11354. .11432
                    /rpt_family="Charlie1a"
     repeat_region  11653. .12821
                    /rpt_family="L2"
     repeat_region  13049. .13281
                    /rpt_family="MIR"
     repeat_region  13369. .13421
                    /rpt_family="Alu"
     repeat_region  13422. .13635
                    /rpt_family="AluSg/x"
     repeat_region  complement(16676. .16878)
                    /rpt_family="L1ME3A"
     STS            16828. .17020
                    /standard_name="149057"
     repeat_region  16879. .16954
                    /rpt_family="(CTTG)n"
     repeat_region  complement(16961. .17268)
                    /rpt_family="AluJo"
     repeat_region  complement(17269. .17288)
                    /rpt_family="L1ME3A"
     repeat_region  17686. .18091
                    /rpt_family="MSTA"
     repeat_region  18096. .19753
                    /rpt_family="MSTA-internal"
     repeat_region  19754. .20163
                    /rpt_family="MSTA"
     repeat_region  21562. .21626
                    /rpt_family="MIR"
     repeat_region  21627. .21671
                    /rpt_family="L2"
     repeat_region  complement(21800. .21901)
                    /rpt_family="L2"
     repeat_region  21902. .22015
                    /rpt_family="L1MC1"
     repeat_region  22016. .22332
```

```
                     /rpt_family="AluSx"
     repeat_region   22333. .22698
                     /rpt_family="L1MC1"
     repeat_region   22699. .22865
                     /rpt_family="MER46A"
     repeat_region   22866. .23466
                     /rpt_family="L1MC1"
     repeat_region   complement(24011. .24096)
                     /rpt_family="MIR"
     repeat_region   complement(24280. .24368)
                     /rpt_family="L1MC4"
     repeat_region   complement(24403. .24678)

  Query Match             58.6%;  Score 17;  DB 8;  Length 157289;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 TGGAGCTCAGGCATGAG 23
              |||||||||||||||||
Db      71025 TGGAGCTCAGGCATGAG 71009


RESULT 43
AC130459
LOCUS       AC130459              157462 bp    DNA     linear   PRI 20-SEP-2002
DEFINITION  Homo sapiens chromosome 16 clone CTA-427H10, complete sequence.
ACCESSION   AC130459
VERSION     AC130459.2  GI:23237942
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 157462)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 157462)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-AUG-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 157462)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-SEP-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Sep 20, 2002 this sequence version replaced gi:22203232.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center and Los Alamos
            National Laboratory
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;
            Estimated Total Number of Errors is 0.2.
FEATURES             Location/Qualifiers
     source          1. .157462
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="CTA-427H10"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 8;  Length 157462;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAGC 24
              |||||||||||||||||

Db      65690 GGAGCTCAGGCATGAGC 65706


RESULT 44
HUAC004626/c
LOCUS       HUAC004626            157838 bp    DNA     linear   PRI 30-OCT-2002
DEFINITION  Homo sapiens Chromosome 16 BAC clone CIT987SK-A-427H10, complete
            sequence.
ACCESSION   AC004626
VERSION     AC004626.1  GI:3337396
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 157838)
  AUTHORS   Loftus,B.J., Kim,U.J., Sneddon,V.P., Kalush,F., Brandon,R.,
            Fuhrmann,J., Mason,T., Crosby,M.L., Barnstead,M., Cronin,L.,
            Deslattes Mays,A., Cao,Y., Xu,R.X., Kang,H.L., Mitchell,S.,
            Eichler,E.E., Harris,P.C., Venter,J.C. and Adams,M.D.
  TITLE     Genome duplications and other features in 12 Mb of DNA sequence
            from human chromosome 16p and 16q
  JOURNAL   Genomics 60 (3), 295-308 (1999)
   PUBMED   10493829
REFERENCE   2  (bases 1 to 157838)
  AUTHORS   Adams,M.D., Loftus,B.J., Zhou,L., Crosby,M., Fuhrmann,J.,
            Mason,T.M., Brandon,R., Kim,U.J., Kerlavage,A.R. and Venter,J.C.
  TITLE     Homo sapiens Chromosome 16 BAC clone CIT987SK-A-427H10
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 157838)
  AUTHORS   Adams,M.D. and Loftus,B.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-APR-1998) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA, Email:
            bjloftus@tigr.org
REFERENCE   4  (bases 1 to 157838)
  AUTHORS   Adams,M.D. and Loftus,B.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JUL-1998) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA
COMMENT     On Jul 24, 1998 this sequence version replaced gi:3212144.
            Address all correspondence to: Mark Adams The Institute for Genomic
            Research 9712 Medical Center Dr, Rockville, MD 20850, USA e-mail
            address: humgen@tigr.org.   The orientation of the sequence is from
            SP6 end to T7 end. Genes were identified by a combination of five
            methods including: XGRAIL (available by anonymous ftp from
            arthur.epm.ornl.gov), Genefinder (Phil Green, University of
            Washington), Genscan  (Chris Burge,
            http://gnomic.stanford.edu/~chris/GENSCANW.html)searches of the
            complete sequence against a peptide database,and the Human gene
            Index database at TIGR (http://www.tigr.org/tdb/hgi/hgi.html).
            Genes without pepetide homolgy having spliced EST hits are termed
            'Unknown gene product'. Genes encoding tRNAs are predicted by
            tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).
FEATURES             Location/Qualifiers
     source          1. .157838
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /map="#16q12.1+16q22/23+1q11/12"
                     /clone="3+1q11/12"
     STS             61033. .61164
                     /note="32826, A002D07, Chr. -, Homo sapiens"
                     /db_xref="dbSTS:G19948"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 8;  Length 157838;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAGC 24
```

```
              ||||||||||||||||||
Db      92149 GGAGCTCAGGCATGAGC 92133


RESULT 45
AC137788
LOCUS       AC137788               158420 bp    DNA     linear   PRI 21-FEB-2003
DEFINITION  Homo sapiens chromosome 16 clone RP11-207M10, complete sequence.
ACCESSION   AC137788
VERSION     AC137788.3  GI:28460742
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 158420)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 158420)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-DEC-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 158420)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 158420)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-FEB-2003) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Feb 21, 2003 this sequence version replaced gi:26023962.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center and Los Alamos
            National Laboratory
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.
FEATURES             Location/Qualifiers
     source          1. .158420
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="RP11-207M10"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 8;  Length 158420;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GGAGCTCAGGCATGAGC 24
              ||||||||||||||||||
Db     101312 GGAGCTCAGGCATGAGC 101328


RESULT 46
AC026533
LOCUS       AC026533               162025 bp    DNA     linear   HTG 26-MAY-2000
DEFINITION  Homo sapiens clone RP11-230J8, WORKING DRAFT SEQUENCE, 34 unordered
            pieces.
ACCESSION   AC026533
VERSION     AC026533.2  GI:8076933
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 162025)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-230J8
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 162025)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On May 25, 2000 this sequence version replaced gi:7283231.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L7054
                Center clone name: 230_J_8
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 143239 bases at least Q40
                Consensus quality: 152967 bases at least Q30
                Consensus quality: 156286 bases at least Q20
                Insert size: 176000; agarose-fp
                Insert size: 158725; sum-of-contigs
                Quality coverage: 3.1 in Q20 bases; agarose-fp
                Quality coverage: 3.4 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 34 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1201: contig of 1201 bp in length
            *     1202     1301: gap of 100 bp
            *     1302     2452: contig of 1151 bp in length
            *     2453     2552: gap of 100 bp
            *     2553     2922: contig of 370 bp in length
            *     2923     3022: gap of 100 bp
            *     3023     4293: contig of 1271 bp in length
            *     4294     4393: gap of 100 bp
            *     4394     6106: contig of 1713 bp in length
```

```
          *     6107     6206: gap of 100 bp
          *     6207     7287: contig of 1081 bp in length
          *     7288     7387: gap of 100 bp
          *     7388     9224: contig of 1837 bp in length
          *     9225     9324: gap of 100 bp
          *     9325    10709: contig of 1385 bp in length
          *    10710    10809: gap of 100 bp
          *    10810    12216: contig of 1407 bp in length
          *    12217    12316: gap of 100 bp
          *    12317    13468: contig of 1152 bp in length
          *    13469    13568: gap of 100 bp
          *    13569    15875: contig of 2307 bp in length
          *    15876    15975: gap of 100 bp
          *    15976    17924: contig of 1949 bp in length
          *    17925    18024: gap of 100 bp
          *    18025    20458: contig of 2434 bp in length
          *    20459    20558: gap of 100 bp
          *    20559    23554: contig of 2996 bp in length
          *    23555    23654: gap of 100 bp
          *    23655    26511: contig of 2857 bp in length
          *    26512    26611: gap of 100 bp
          *    26612    29242: contig of 2631 bp in length
          *    29243    29342: gap of 100 bp
          *    29343    33489: contig of 4147 bp in length
          *    33490    33589: gap of 100 bp
          *    33590    37017: contig of 3428 bp in length
          *    37018    37117: gap of 100 bp
          *    37118    40747: contig of 3630 bp in length
          *    40748    40847: gap of 100 bp
          *    40848    44927: contig of 4080 bp in length
          *    44928    45027: gap of 100 bp
          *    45028    49503: contig of 4476 bp in length
          *    49504    49603: gap of 100 bp
          *    49604    55602: contig of 5999 bp in length
          *    55603    55702: gap of 100 bp
          *    55703    59835: contig of 4133 bp in length
          *    59836    59935: gap of 100 bp
          *    59936    64812: contig of 4877 bp in length
          *    64813    64912: gap of 100 bp
          *    64913    71788: contig of 6876 bp in length
          *    71789    71888: gap of 100 bp
          *    71889    78138: contig of 6250 bp in length
          *    78139    78238: gap of 100 bp
          *    78239    84840: contig of 6602 bp in length
          *    84841    84940: gap of 100 bp
          *    84941    93044: contig of 8104 bp in length
          *    93045    93144: gap of 100 bp
          *    93145   100167: contig of 7023 bp in length
          *   100168   100267: gap of 100 bp
          *   100268   110560: contig of 10293 bp in length
          *   110561   110660: gap of 100 bp
          *   110661   122758: contig of 12098 bp in length
          *   122759   122858: gap of 100 bp
          *   122859   131690: contig of 8832 bp in length
          *   131691   131790: gap of 100 bp
          *   131791   147247: contig of 15457 bp in length
          *   147248   147347: gap of 100 bp
          *   147348   162025: contig of 14678 bp in length.
FEATURES             Location/Qualifiers
     source          1. .162025
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-230J8"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1201
                     /note="assembly_fragment"
     gap             1202. .1301
                     /estimated_length=100
     misc_feature    1302. .2452
                     /note="assembly_fragment"
     gap             2453. .2552
                     /estimated_length=100
     misc_feature    2553. .2922
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     gap             2923. .3022
                     /estimated_length=100
     misc_feature    3023. .4293
                     /note="assembly_fragment"
     gap             4294. .4393
                     /estimated_length=100
     misc_feature    4394. .6106
                     /note="assembly_fragment"
     gap             6107. .6206
                     /estimated_length=100
     misc_feature    6207. .7287
                     /note="assembly_fragment"
     gap             7288. .7387
                     /estimated_length=100
     misc_feature    7388. .9224
                     /note="assembly_fragment"
     gap             9225. .9324
                     /estimated_length=100
     misc_feature    9325. .10709
                     /note="assembly_fragment"
     gap             10710. .10809
                     /estimated_length=100
     misc_feature    10810. .12216
                     /note="assembly_fragment"
     gap             12217. .12316
                     /estimated_length=100
     misc_feature    12317. .13468
                     /note="assembly_fragment"
     gap             13469. .13568
                     /estimated_length=100
     misc_feature    13569. .15875
                     /note="assembly_fragment"
     gap             15876. .15975
                     /estimated_length=100
     misc_feature    15976. .17924
                     /note="assembly_fragment"
     gap             17925. .18024
                     /estimated_length=100
     misc_feature    18025. .20458
                     /note="assembly_fragment"
     gap             20459. .20558
                     /estimated_length=100

  Query Match             58.6%;  Score 17;  DB 14;  Length 162025;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          13 TCAGGCATGAGCCAGCA 29
               |||||||||||||||||
Db       31714 TCAGGCATGAGCCAGCA 31730


RESULT 47
AC164010/c
LOCUS       AC164010              162585 bp    DNA     linear   ROD 10-AUG-2005
DEFINITION  Mus musculus chromosome 18, clone RP24-391C15, complete sequence.
ACCESSION   AC164010
VERSION     AC164010.2  GI:72097583
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 162585)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus chromosome 18, clone RP24-391C15
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 162585)
```

```
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
            Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
            Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
            DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
            Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
            Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
            Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
            Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
            MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
            Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
            Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2005) Broad Institute of MIT and Harvard, 320
            Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 162585)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
            Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
            Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
            DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
            Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
            Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
            Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
            Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
            MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
            Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
            Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-AUG-2005) Broad Institute of MIT and Harvard, 320
            Charles Street, Cambridge, MA 02141, USA
COMMENT     On Aug 10, 2005 this sequence version replaced gi:67764034.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Broad Institute of MIT and Harvard
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@broad.mit.edu
            -------------- Project Information
                Center project name: J5489
                Center clone name: 391_C_15
            --------------

            Some of the sequence contained within base pairs 1 - 98176 was
            stolen from accession AC109243.
FEATURES              Location/Qualifiers
     source           1. .162585
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:10090"
                      /chromosome="18"
                      /map="18"
                      /clone="RP24-391C15"
                      /clone_lib="RPCI-24 Male Mouse BAC"
     repeat_region    complement(2372. .2529)
                      /rpt_family="MER5A"
     repeat_region    3729. .4005
                      /rpt_family="ORR1D1"
     repeat_region    complement(4340. .4651)
                      /rpt_family="RLTR28"
     repeat_region    complement(4652. .11218)
                      /rpt_family="L1Md_T"
     repeat_region    complement(11219. .11388)
                      /rpt_family="RLTR28"
     repeat_region    complement(14200. .14296)
                      /rpt_family="MIRb"
     repeat_region    complement(15802. .15908)
                      /rpt_family="MIRb"
     repeat_region    complement(15927. .16086)
                      /rpt_family="Lx"
     repeat_region    17097. .17182
                      /rpt_family="(TC)n"
     repeat_region    17309. .17361
                      /rpt_family="(TG)n"
     repeat_region    17498. .17544
                      /rpt_family="L3"
     repeat_region    20748. .20873
                      /rpt_family="(TG)n"
     repeat_region    complement(22782. .23122)
                      /rpt_family="ORR1E"
     repeat_region    complement(23168. .23534)
                      /rpt_family="MTD"
     repeat_region    complement(24103. .24315)
                      /rpt_family="Charlie5"
     repeat_region    complement(24324. .24430)
                      /rpt_family="MER33"
     repeat_region    complement(24562. .24656)
                      /rpt_family="ID_B1"
     repeat_region    25624. .25660
                      /rpt_family="(T)n"
     repeat_region    complement(27625. .27747)
                      /rpt_family="B2_Mm1t"
     repeat_region    27853. .27926
                      /rpt_family="MIR"
     repeat_region    30096. .30181
                      /rpt_family="MTD"
     repeat_region    30172. .30350
                      /rpt_family="MTD"
     repeat_region    30851. .30906
                      /rpt_family="(TA)n"
     repeat_region    complement(31586. .32328)
                      /rpt_family="RMER19A"
     repeat_region    32476. .32584
                      /rpt_family="PB1D10"
     repeat_region    complement(32628. .32925)
                      /rpt_family="L1MC4a"
     repeat_region    34200. .34224
                      /rpt_family="(TG)n"
     repeat_region    34303. .34450
                      /rpt_family="B1F2"
     repeat_region    complement(34686. .34819)
                      /rpt_family="MLT1A0"
     repeat_region    34820. .34966
                      /rpt_family="B1_Mm"
     repeat_region    35046. .35083
                      /rpt_family="L2"
     repeat_region    35281. .35470
                      /rpt_family="MIR"
     repeat_region    36758. .36899
                      /rpt_family="MER5B"
     repeat_region    37776. .37853
                      /rpt_family="(TG)n"
     repeat_region    37904. .37934
                      /rpt_family="(TG)n"
     repeat_region    37963. .38048
                      /rpt_family="(TG)n"
```

```
     repeat_region   38077. .38121
                     /rpt_family="(TG)n"
     repeat_region   38968. .39009
                     /rpt_family="(TG)n"
     repeat_region   complement(39707. .39841)
                     /rpt_family="MIRb"
     repeat_region   40048. .40096
                     /rpt_family="(TC)n"
     repeat_region   complement(40098. .40207)
                     /rpt_family="ID_B1"
     repeat_region   43619. .43941
                     /rpt_family="MTD"
     repeat_region   44110. .44135
                     /rpt_family="(CCCCCA)n"
     repeat_region   44143. .44286
                     /rpt_family="RSINE1"
     repeat_region   44287. .44306
                     /rpt_family="(CA)n"
     repeat_region   45151. .45336
                     /rpt_family="B2_Mm2"
     repeat_region   46296. .46342
                     /rpt_family="B1F"
     repeat_region   46632. .46750
                     /rpt_family="B1_Mur4"
     repeat_region   47535. .47641
                     /rpt_family="L1MB3"
     repeat_region   47714. .47908
                     /rpt_family="B3"
     repeat_region   47923. .47962
                     /rpt_family="(GA)n"
     repeat_region   47980. .48137
                     /rpt_family="B3A"
     repeat_region   51319. .51364
                     /rpt_family="(GA)n"
     repeat_region   52604. .52649
                     /rpt_family="T-rich"
     repeat_region   52952. .53040
                     /rpt_family="(CATATA)n"
     repeat_region   complement(53044. .53127)

  Query Match             58.6%;  Score 17;  DB 9;  Length 162585;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CTCTCTGGAGCTCAGGC 18
              |||||||||||||||||
Db      23160 CTCTCTGGAGCTCAGGC 23144


RESULT 48
AL808118
LOCUS       AL808118              164068 bp    DNA     linear   ROD 07-NOV-2002
DEFINITION  Mouse DNA sequence from clone RP23-287O8 on chromosome 4, complete
            sequence.
ACCESSION   AL808118
VERSION     AL808118.8  GI:24366576
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 164068)
  AUTHORS   Johnson,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 24, 2002 this sequence version replaced gi:23895253.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-287O8 is
            from the RPCI-23 Mouse PAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .164068
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="4"
                     /clone="RP23-287O8"
                     /clone_lib="RPCI-23"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 9;  Length 164068;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 CTCTCTGGAGCTCAGGC 18
              |||||||||||||||||
Db      69589 CTCTCTGGAGCTCAGGC 69605


RESULT 49
AL451066
LOCUS       AL451066              166258 bp    DNA     linear   HTG 29-JAN-2004
DEFINITION  Homo sapiens chromosome 1 clone RP11-262P12, 11 unordered pieces.
ACCESSION   AL451066 AC031996
VERSION     AL451066.9  GI:14586202
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 166258)
  AUTHORS   Almeida,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JAN-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jul 2, 2001 this sequence version replaced gi:14529893.
            Draft Sequence Produced by Whitehead Institute/MIT Center for
            Genome Research, 320 Charles Street,
            Cambridge, MA 02141, USA
            http://www-seq.wi.mit.edu
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bA262P12
```

```
            ------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: M13; M77815; 31% of reads
            Sequencing vector: plasmid; L08752; 68% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 162926 bases at least Q40
            Consensus quality: 163856 bases at least Q30
            Consensus quality: 164540 bases at least Q20
            Insert size: 165258; sum-of-contigs
            Insert size: 165960; 2.4% error; agarose-fp
            Quality coverage: 6.74x in Q20 bases; sum-of-contigs Quality
            coverage: 8.02x in Q20 bases; agarose-fp
            -------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 11 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     5862: contig of 5862 bp in length
            *    5863     5962: gap of 100 bp
            *    5963    14007: contig of 8045 bp in length
            *   14008    14107: gap of 100 bp
            *   14108    30475: contig of 16368 bp in length
            *   30476    30575: gap of 100 bp
            *   30576    55381: contig of 24806 bp in length
            *   55382    55481: gap of 100 bp
            *   55482    90435: contig of 34954 bp in length
            *   90436    90535: gap of 100 bp
            *   90536   112697: contig of 22162 bp in length
            *  112698   112797: gap of 100 bp
            *  112798   118156: contig of 5359 bp in length
            *  118157   118256: gap of 100 bp
            *  118257   150482: contig of 32226 bp in length
            *  150483   150582: gap of 100 bp
            *  150583   154622: contig of 4040 bp in length
            *  154623   154722: gap of 100 bp
            *  154723   157118: contig of 2396 bp in length
            *  157119   157218: gap of 100 bp
            *  157219   166258: contig of 9040 bp in length.
FEATURES             Location/Qualifiers
     source          1. .166258
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-262P12"
                     /clone_lib="RPCI-11.1"
     misc_feature    1. .5862
                     /note="assembly_fragment:02667
                     fragment_chain:1
                     clone_end:SP6
                     vector_side:left"
     misc_feature    5963. .14007
                     /note="assembly_fragment:00792
                     fragment_chain:1"
     misc_feature    14108. .30475
                     /note="assembly_fragment:02400
                     fragment_chain:1"
     misc_feature    30576. .55381
                     /note="assembly_fragment:02951
                     fragment_chain:1"
     misc_feature    55482. .90435
                     /note="assembly_fragment:00679
                     fragment_chain:1"
     misc_feature    90536. .112697
                     /note="assembly_fragment:03378
                     fragment_chain:1"
     misc_feature    112798. .118156
                     /note="assembly_fragment:03132
                     fragment_chain:1"
     misc_feature    118257. .150482
                     /note="assembly_fragment:02735
                     fragment_chain:1"
     misc_feature    150583. .154622
                     /note="assembly_fragment:00096"
     misc_feature    154723. .157118
                     /note="assembly_fragment:00612"
     misc_feature    157219. .166258
                     /note="assembly_fragment:00806
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 14;  Length 166258;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db      12620 TCAGGCATGAGCCAGCA 12636


RESULT 50
AC020990
LOCUS       AC020990              166418 bp    DNA     linear   HTG 16-MAR-2000
DEFINITION  Homo sapiens clone RP11-20N5, WORKING DRAFT SEQUENCE, 11 unordered
            pieces.
ACCESSION   AC020990
VERSION     AC020990.4  GI:7249019
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 166418)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-20N5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 166418)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 16, 2000 this sequence version replaced gi:6778547.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L3946
                Center clone name: 20_N_5
            -------------- Summary Statistics
```

```
            Sequencing vector: M13; M77815; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.960731
            Consensus quality: 160812 bases at least Q40
            Consensus quality: 163732 bases at least Q30
            Consensus quality: 164645 bases at least Q20
            Insert size: 165000; agarose-fp
            Insert size: 165418; sum-of-contigs
            Quality coverage: 4.9 in Q20 bases; agarose-fp
            Quality coverage: 4.9 in Q20 bases; sum-of-contigs
          --------------.
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 11 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *         1      345: contig of 345 bp in length
          *       346      445: gap of 100 bp
          *       446     2083: contig of 1638 bp in length
          *      2084     2183: gap of 100 bp
          *      2184     5811: contig of 3628 bp in length
          *      5812     5911: gap of 100 bp
          *      5912    12018: contig of 6107 bp in length
          *     12019    12118: gap of 100 bp
          *     12119    18513: contig of 6395 bp in length
          *     18514    18613: gap of 100 bp
          *     18614    25785: contig of 7172 bp in length
          *     25786    25885: gap of 100 bp
          *     25886    36446: contig of 10561 bp in length
          *     36447    36546: gap of 100 bp
          *     36547    59979: contig of 23433 bp in length
          *     59980    60079: gap of 100 bp
          *     60080    93312: contig of 33233 bp in length
          *     93313    93412: gap of 100 bp
          *     93413   128579: contig of 35167 bp in length
          *    128580   128679: gap of 100 bp
          *    128680   166418: contig of 37739 bp in length.
FEATURES             Location/Qualifiers
     source          1. .166418
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-20N5"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .345
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     gap             346. .445
                     /estimated_length=100
     misc_feature    446. .2083
                     /note="assembly_fragment"
     gap             2084. .2183
                     /estimated_length=100
     misc_feature    2184. .5811
                     /note="assembly_fragment"
     gap             5812. .5911
                     /estimated_length=100
     misc_feature    5912. .12018
                     /note="assembly_fragment"
     gap             12019. .12118
                     /estimated_length=100
     misc_feature    12119. .18513
                     /note="assembly_fragment"
     gap             18514. .18613
                     /estimated_length=100
     misc_feature    18614. .25785
                     /note="assembly_fragment"
     gap             25786. .25885
                     /estimated_length=100
     misc_feature    25886. .36446
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     gap             36447. .36546
                     /estimated_length=100
     misc_feature    36547. .59979
                     /note="assembly_fragment"
     gap             59980. .60079
                     /estimated_length=100
     misc_feature    60080. .93312
                     /note="assembly_fragment"
     gap             93313. .93412
                     /estimated_length=100
     misc_feature    93413. .128579
                     /note="assembly_fragment"
     gap             128580. .128679
                     /estimated_length=100
     misc_feature    128680. .166418
                     /note="assembly_fragment"
ORIGIN

  Query Match             58.6%;  Score 17;  DB 14;  Length 166418;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 GAGCTCAGGCATGAGCC 25
              |||||||||||||||||
Db      17922 GAGCTCAGGCATGAGCC 17938


Search completed: April 12, 2006, 15:04:58
Job time : 1866 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: April 12, 2006, 14:01:46 ; Search time 680 Seconds
(without alignments)
171.775 Million cell updates/sec

Title: SEQ1-4023-4051-4037A
Perfect score: 29
Sequence: 1 cctctctggagctcaggcatgagccagca 29

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 9281099 seqs, 2013915447 residues

Word size : 15

Total number of hits satisfying chosen parameters: 521

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Published_Applications_NA_New:*
1: /SIDS5/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /SIDS5/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /SIDS5/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /SIDS5/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /SIDS5/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /SIDS5/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /SIDS5/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /SIDS5/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /SIDS5/ptodata/2/pubpna/US10_NEW_PUB.seq2:*
10: /SIDS5/ptodata/2/pubpna/US10_NEW_PUB.seq3:*
11: /SIDS5/ptodata/2/pubpna/US11_NEW_PUB.seq:*
12: /SIDS5/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
13: /SIDS5/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
14: /SIDS5/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
15: /SIDS5/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	17	58.6	634	6	US-09-925-065A-577180	Sequence 577180,
c	2	17	58.6	634	6	US-09-925-065A-577181	Sequence 577181,
c	3	17	58.6	634	6	US-09-925-065A-577182	Sequence 577182,
c	4	17	58.6	21728	9	US-10-330-773-362	Sequence 362, App
	5	16	55.2	22	8	US-10-310-914A-1280357	Sequence 1280357,
	6	16	55.2	22	8	US-10-310-914A-1280364	Sequence 1280364,
	7	16	55.2	24	8	US-10-310-914A-1280360	Sequence 1280360,
	8	16	55.2	24	8	US-10-310-914A-1280389	Sequence 1280389,
	9	16	55.2	25	8	US-10-310-914A-1280390	Sequence 1280390,
	10	16	55.2	25	14	US-11-136-527-241234	Sequence 241234,
	11	16	55.2	25	14	US-11-136-527-241243	Sequence 241243,
	12	16	55.2	25	14	US-11-136-527-241246	Sequence 241246,
	13	16	55.2	27	8	US-10-310-914A-1280391	Sequence 1280391,
c	14	16	55.2	201	8	US-10-995-561-41746	Sequence 41746, A
c	15	16	55.2	201	8	US-10-995-561-79446	Sequence 79446, A
c	16	16	55.2	201	8	US-10-995-561-79792	Sequence 79792, A
	17	16	55.2	201	14	US-11-124-367A-15249	Sequence 15249, A
c	18	16	55.2	201	14	US-11-124-367A-29976	Sequence 29976, A
	19	16	55.2	293	9	US-10-301-480-207048	Sequence 207048,
	20	16	55.2	293	9	US-10-301-480-207050	Sequence 207050,
	21	16	55.2	293	10	US-10-301-480-820457	Sequence 820457,
	22	16	55.2	293	10	US-10-301-480-820459	Sequence 820459,
	23	16	55.2	299	6	US-09-925-065A-107140	Sequence 107140,
	24	16	55.2	299	6	US-09-925-065A-107142	Sequence 107142,
c	25	16	55.2	370	6	US-09-925-065A-431188	Sequence 431188,
	26	16	55.2	373	10	US-10-301-480-493877	Sequence 493877,
	27	16	55.2	373	10	US-10-301-480-1107286	Sequence 1107286,
c	28	16	55.2	383	10	US-10-301-480-252290	Sequence 252290,
c	29	16	55.2	383	10	US-10-301-480-865699	Sequence 865699,
c	30	16	55.2	384	10	US-10-301-480-252286	Sequence 252286,
c	31	16	55.2	384	10	US-10-301-480-865695	Sequence 865695,
c	32	16	55.2	386	6	US-09-925-065A-159403	Sequence 159403,
c	33	16	55.2	387	6	US-09-925-065A-159399	Sequence 159399,
c	34	16	55.2	388	9	US-10-301-480-204513	Sequence 204513,
c	35	16	55.2	388	10	US-10-301-480-252293	Sequence 252293,
c	36	16	55.2	388	10	US-10-301-480-817922	Sequence 817922,
c	37	16	55.2	388	10	US-10-301-480-865702	Sequence 865702,
c	38	16	55.2	391	6	US-09-925-065A-104346	Sequence 104346,
c	39	16	55.2	391	6	US-09-925-065A-159406	Sequence 159406,
c	40	16	55.2	392	10	US-10-301-480-252287	Sequence 252287,
c	41	16	55.2	392	10	US-10-301-480-865696	Sequence 865696,
c	42	16	55.2	395	6	US-09-925-065A-159400	Sequence 159400,
c	43	16	55.2	396	6	US-09-925-065A-173155	Sequence 173155,
c	44	16	55.2	396	10	US-10-301-480-264760	Sequence 264760,
c	45	16	55.2	396	10	US-10-301-480-878169	Sequence 878169,
c	46	16	55.2	400	6	US-09-925-065A-173157	Sequence 173157,
c	47	16	55.2	400	6	US-09-925-065A-173158	Sequence 173158,
c	48	16	55.2	400	6	US-09-925-065A-173159	Sequence 173159,
c	49	16	55.2	400	9	US-10-301-480-204515	Sequence 204515,
c	50	16	55.2	400	10	US-10-301-480-264762	Sequence 264762,
c	51	16	55.2	400	10	US-10-301-480-264763	Sequence 264763,
c	52	16	55.2	400	10	US-10-301-480-264764	Sequence 264764,
c	53	16	55.2	400	10	US-10-301-480-817924	Sequence 817924,
c	54	16	55.2	400	10	US-10-301-480-878171	Sequence 878171,
c	55	16	55.2	400	10	US-10-301-480-878172	Sequence 878172,
c	56	16	55.2	400	10	US-10-301-480-878173	Sequence 878173,
c	57	16	55.2	401	6	US-09-925-065A-104348	Sequence 104348,
c	58	16	55.2	401	6	US-09-925-065A-159410	Sequence 159410,
	59	16	55.2	404	9	US-10-301-480-207047	Sequence 207047,
	60	16	55.2	404	10	US-10-301-480-261341	Sequence 261341,
	61	16	55.2	404	10	US-10-301-480-820456	Sequence 820456,
	62	16	55.2	404	10	US-10-301-480-874750	Sequence 874750,
	63	16	55.2	420	6	US-09-925-065A-107139	Sequence 107139,
	64	16	55.2	420	6	US-09-925-065A-169368	Sequence 169368,
	65	16	55.2	420	6	US-09-925-065A-169369	Sequence 169369,
c	66	16	55.2	420	10	US-10-301-480-252291	Sequence 252291,
c	67	16	55.2	420	10	US-10-301-480-865700	Sequence 865700,
c	68	16	55.2	421	9	US-10-301-480-204514	Sequence 204514,
c	69	16	55.2	421	10	US-10-301-480-817923	Sequence 817923,
c	70	16	55.2	423	6	US-09-925-065A-159404	Sequence 159404,
c	71	16	55.2	423	10	US-10-301-480-252288	Sequence 252288,
c	72	16	55.2	423	10	US-10-301-480-252292	Sequence 252292,
c	73	16	55.2	423	10	US-10-301-480-252294	Sequence 252294,
c	74	16	55.2	423	10	US-10-301-480-252295	Sequence 252295,
c	75	16	55.2	423	10	US-10-301-480-252296	Sequence 252296,
c	76	16	55.2	423	10	US-10-301-480-865697	Sequence 865697,
c	77	16	55.2	423	10	US-10-301-480-865701	Sequence 865701,
c	78	16	55.2	423	10	US-10-301-480-865703	Sequence 865703,
c	79	16	55.2	423	10	US-10-301-480-865704	Sequence 865704,
c	80	16	55.2	423	10	US-10-301-480-865705	Sequence 865705,
c	81	16	55.2	424	6	US-09-925-065A-104347	Sequence 104347,
c	82	16	55.2	426	6	US-09-925-065A-159401	Sequence 159401,
c	83	16	55.2	426	6	US-09-925-065A-159405	Sequence 159405,
c	84	16	55.2	426	6	US-09-925-065A-159407	Sequence 159407,
c	85	16	55.2	426	6	US-09-925-065A-159408	Sequence 159408,
c	86	16	55.2	426	6	US-09-925-065A-159409	Sequence 159409,
	87	16	55.2	462	6	US-09-925-065A-483021	Sequence 483021,
c	88	16	55.2	491	6	US-09-925-065A-30838	Sequence 30838, A
c	89	16	55.2	491	9	US-10-301-480-132076	Sequence 132076,
c	90	16	55.2	491	10	US-10-301-480-745485	Sequence 745485,
c	91	16	55.2	501	6	US-09-925-065A-577939	Sequence 577939,

c 92	16	55.2	520	6	US-09-925-065A-21217	Sequence 21217, A
c 93	16	55.2	520	9	US-10-301-480-122454	Sequence 122454,
c 94	16	55.2	520	10	US-10-301-480-735863	Sequence 735863,
95	16	55.2	529	6	US-09-925-065A-444539	Sequence 444539,
96	16	55.2	532	6	US-09-925-065A-745207	Sequence 745207,
97	16	55.2	532	6	US-09-925-065A-745208	Sequence 745208,
c 98	16	55.2	533	6	US-09-925-065A-476643	Sequence 476643,
c 99	16	55.2	533	6	US-09-925-065A-476644	Sequence 476644,
100	16	55.2	536	10	US-10-301-480-504776	Sequence 504776,
101	16	55.2	536	10	US-10-301-480-1118185	Sequence 1118185,
102	16	55.2	541	6	US-09-925-065A-740567	Sequence 740567,
c 103	16	55.2	544	10	US-10-301-480-548890	Sequence 548890,
c 104	16	55.2	544	10	US-10-301-480-1162299	Sequence 1162299,
105	16	55.2	548	6	US-09-925-065A-29550	Sequence 29550, A
106	16	55.2	548	9	US-10-301-480-130788	Sequence 130788,
107	16	55.2	548	10	US-10-301-480-744197	Sequence 744197,
108	16	55.2	549	6	US-09-925-065A-556023	Sequence 556023,
109	16	55.2	549	6	US-09-925-065A-556024	Sequence 556024,
110	16	55.2	549	6	US-09-925-065A-841088	Sequence 841088,
111	16	55.2	551	6	US-09-925-065A-928034	Sequence 928034,
112	16	55.2	551	6	US-09-925-065A-930207	Sequence 930207,
113	16	55.2	565	6	US-09-925-065A-777985	Sequence 777985,
c 114	16	55.2	567	6	US-09-925-065A-556758	Sequence 556758,
c 115	16	55.2	567	6	US-09-925-065A-656344	Sequence 656344,
c 116	16	55.2	567	10	US-10-301-480-236486	Sequence 236486,
c 117	16	55.2	567	10	US-10-301-480-236487	Sequence 236487,
c 118	16	55.2	567	10	US-10-301-480-849895	Sequence 849895,
c 119	16	55.2	567	10	US-10-301-480-849896	Sequence 849896,
c 120	16	55.2	568	6	US-09-925-065A-141620	Sequence 141620,
c 121	16	55.2	568	6	US-09-925-065A-141621	Sequence 141621,
122	16	55.2	574	6	US-09-925-065A-799743	Sequence 799743,
c 123	16	55.2	579	10	US-10-301-480-435946	Sequence 435946,
c 124	16	55.2	579	10	US-10-301-480-1049355	Sequence 1049355,
c 125	16	55.2	585	6	US-09-925-065A-365382	Sequence 365382,
126	16	55.2	585	6	US-09-925-065A-892775	Sequence 892775,
c 127	16	55.2	590	10	US-10-301-480-340191	Sequence 340191,
c 128	16	55.2	590	10	US-10-301-480-340192	Sequence 340192,
c 129	16	55.2	590	10	US-10-301-480-953600	Sequence 953600,
c 130	16	55.2	590	10	US-10-301-480-953601	Sequence 953601,
131	16	55.2	591	9	US-10-301-480-100412	Sequence 100412,
132	16	55.2	591	10	US-10-301-480-713821	Sequence 713821,
c 133	16	55.2	593	6	US-09-925-065A-899570	Sequence 899570,
c 134	16	55.2	594	6	US-09-925-065A-261937	Sequence 261937,
c 135	16	55.2	594	6	US-09-925-065A-261938	Sequence 261938,
136	16	55.2	594	6	US-09-925-065A-519228	Sequence 519228,
137	16	55.2	595	6	US-09-925-065A-483020	Sequence 483020,
138	16	55.2	595	6	US-09-925-065A-483022	Sequence 483022,
139	16	55.2	598	6	US-09-925-065A-869776	Sequence 869776,
140	16	55.2	600	14	US-11-136-527-5111	Sequence 5111, Ap
141	16	55.2	606	6	US-09-925-065A-680666	Sequence 680666,
c 142	16	55.2	607	6	US-09-925-065A-737401	Sequence 737401,
c 143	16	55.2	607	10	US-10-301-480-439980	Sequence 439980,
c 144	16	55.2	607	10	US-10-301-480-439981	Sequence 439981,
c 145	16	55.2	607	10	US-10-301-480-1053389	Sequence 1053389,
c 146	16	55.2	607	10	US-10-301-480-1053390	Sequence 1053390,
147	16	55.2	610	6	US-09-925-065A-767246	Sequence 767246,
c 148	16	55.2	621	6	US-09-925-065A-765518	Sequence 765518,
c 149	16	55.2	622	6	US-09-925-065A-369908	Sequence 369908,
c 150	16	55.2	622	6	US-09-925-065A-369909	Sequence 369909,
151	16	55.2	627	6	US-09-925-065A-560490	Sequence 560490,
c 152	16	55.2	629	6	US-09-925-065A-252792	Sequence 252792,
c 153	16	55.2	629	6	US-09-925-065A-252793	Sequence 252793,
c 154	16	55.2	629	6	US-09-925-065A-252794	Sequence 252794,
c 155	16	55.2	629	6	US-09-925-065A-252795	Sequence 252795,
c 156	16	55.2	632	10	US-10-301-480-332197	Sequence 332197,
c 157	16	55.2	632	10	US-10-301-480-332198	Sequence 332198,
c 158	16	55.2	632	10	US-10-301-480-332199	Sequence 332199,
c 159	16	55.2	632	10	US-10-301-480-332200	Sequence 332200,
c 160	16	55.2	632	10	US-10-301-480-945606	Sequence 945606,
c 161	16	55.2	632	10	US-10-301-480-945607	Sequence 945607,
c 162	16	55.2	632	10	US-10-301-480-945608	Sequence 945608,
c 163	16	55.2	632	10	US-10-301-480-945609	Sequence 945609,
164	16	55.2	635	6	US-09-925-065A-84844	Sequence 84844, A
165	16	55.2	635	9	US-10-301-480-186084	Sequence 186084,
166	16	55.2	635	10	US-10-301-480-799493	Sequence 799493,
167	16	55.2	638	6	US-09-925-065A-800541	Sequence 800541,
168	16	55.2	650	10	US-10-301-480-602622	Sequence 602622,
169	16	55.2	650	10	US-10-301-480-1216031	Sequence 1216031,
c 170	16	55.2	655	6	US-09-925-065A-716468	Sequence 716468,
171	16	55.2	657	6	US-09-925-065A-955783	Sequence 955783,
172	16	55.2	659	6	US-09-925-065A-767245	Sequence 767245,
173	16	55.2	659	6	US-09-925-065A-767247	Sequence 767247,
c 174	16	55.2	668	9	US-10-301-480-58856	Sequence 58856, A
c 175	16	55.2	668	9	US-10-301-480-58857	Sequence 58857, A
c 176	16	55.2	668	10	US-10-301-480-672265	Sequence 672265,
c 177	16	55.2	668	10	US-10-301-480-672266	Sequence 672266,
178	16	55.2	674	6	US-09-925-065A-767248	Sequence 767248,
179	16	55.2	693	6	US-09-925-065A-55079	Sequence 55079, A
180	16	55.2	693	6	US-09-925-065A-55080	Sequence 55080, A
181	16	55.2	693	9	US-10-301-480-156317	Sequence 156317,
182	16	55.2	693	9	US-10-301-480-156318	Sequence 156318,
183	16	55.2	693	10	US-10-301-480-769726	Sequence 769726,
184	16	55.2	693	10	US-10-301-480-769727	Sequence 769727,
185	16	55.2	709	10	US-10-301-480-574113	Sequence 574113,
186	16	55.2	709	10	US-10-301-480-1187522	Sequence 1187522,
187	16	55.2	745	6	US-09-925-065A-925382	Sequence 925382,
188	16	55.2	869	10	US-10-301-480-588217	Sequence 588217,
189	16	55.2	869	10	US-10-301-480-1201626	Sequence 1201626,
c 190	16	55.2	875	6	US-09-925-065A-9533	Sequence 9533, Ap
c 191	16	55.2	875	9	US-10-301-480-110770	Sequence 110770,
c 192	16	55.2	875	10	US-10-301-480-724179	Sequence 724179,
c 193	16	55.2	954	10	US-10-301-480-553174	Sequence 553174,
c 194	16	55.2	954	10	US-10-301-480-1166583	Sequence 1166583,
c 195	16	55.2	972	10	US-10-301-480-577723	Sequence 577723,
c 196	16	55.2	972	10	US-10-301-480-577724	Sequence 577724,
c 197	16	55.2	972	10	US-10-301-480-1191132	Sequence 1191132,
c 198	16	55.2	972	10	US-10-301-480-1191133	Sequence 1191133,
c 199	16	55.2	975	10	US-10-301-480-577722	Sequence 577722,
c 200	16	55.2	975	10	US-10-301-480-1191131	Sequence 1191131,
c 201	16	55.2	976	10	US-10-301-480-577720	Sequence 577720,
c 202	16	55.2	976	10	US-10-301-480-577721	Sequence 577721,
c 203	16	55.2	976	10	US-10-301-480-1191129	Sequence 1191129,
c 204	16	55.2	976	10	US-10-301-480-1191130	Sequence 1191130,
205	16	55.2	987	10	US-10-301-480-571139	Sequence 571139,
206	16	55.2	987	10	US-10-301-480-1184548	Sequence 1184548,
207	16	55.2	989	10	US-10-301-480-571140	Sequence 571140,
208	16	55.2	989	10	US-10-301-480-1184549	Sequence 1184549,
c 209	16	55.2	997	10	US-10-301-480-570912	Sequence 570912,
210	16	55.2	997	10	US-10-301-480-571586	Sequence 571586,
211	16	55.2	997	10	US-10-301-480-571587	Sequence 571587,
c 212	16	55.2	997	10	US-10-301-480-1184321	Sequence 1184321,
213	16	55.2	997	10	US-10-301-480-1184995	Sequence 1184995,
214	16	55.2	997	10	US-10-301-480-1184996	Sequence 1184996,
215	16	55.2	999	10	US-10-301-480-533197	Sequence 533197,
216	16	55.2	999	10	US-10-301-480-571573	Sequence 571573,
217	16	55.2	999	10	US-10-301-480-571574	Sequence 571574,
218	16	55.2	999	10	US-10-301-480-1146606	Sequence 1146606,
219	16	55.2	999	10	US-10-301-480-1184982	Sequence 1184982,
220	16	55.2	999	10	US-10-301-480-1184983	Sequence 1184983,
221	16	55.2	1044	6	US-09-925-065A-74190	Sequence 74190, A
222	16	55.2	1044	9	US-10-301-480-175429	Sequence 175429,
223	16	55.2	1044	10	US-10-301-480-788838	Sequence 788838,
c 224	16	55.2	1073	6	US-09-925-065A-697158	Sequence 697158,
c 225	16	55.2	1087	6	US-09-925-065A-678822	Sequence 678822,
c 226	16	55.2	1087	6	US-09-925-065A-678823	Sequence 678823,
c 227	16	55.2	1087	6	US-09-925-065A-678824	Sequence 678824,
c 228	16	55.2	1087	6	US-09-925-065A-678825	Sequence 678825,
229	16	55.2	1148	6	US-09-925-065A-727806	Sequence 727806,
230	16	55.2	1311	6	US-09-925-065A-704573	Sequence 704573,
231	16	55.2	1327	9	US-10-301-480-38989	Sequence 38989, A
232	16	55.2	1327	9	US-10-301-480-38990	Sequence 38990, A
233	16	55.2	1327	9	US-10-301-480-38991	Sequence 38991, A
234	16	55.2	1327	9	US-10-301-480-38992	Sequence 38992, A
235	16	55.2	1327	10	US-10-301-480-652398	Sequence 652398,
236	16	55.2	1327	10	US-10-301-480-652399	Sequence 652399,
237	16	55.2	1327	10	US-10-301-480-652400	Sequence 652400,

	238	16	55.2	1327	10	US-10-301-480-652401 Sequence 652401,
c	239	16	55.2	1375	6	US-09-925-065A-79996 Sequence 79996, A
c	240	16	55.2	1375	6	US-09-925-065A-79997 Sequence 79997, A
c	241	16	55.2	1375	9	US-10-301-480-181235 Sequence 181235,
c	242	16	55.2	1375	9	US-10-301-480-181236 Sequence 181236,
c	243	16	55.2	1375	10	US-10-301-480-794644 Sequence 794644,
c	244	16	55.2	1375	10	US-10-301-480-794645 Sequence 794645,
	245	16	55.2	1399	6	US-09-925-065A-717278 Sequence 717278,
c	246	16	55.2	1729	6	US-09-925-065A-714716 Sequence 714716,
c	247	16	55.2	1866	8	US-10-750-185-45574 Sequence 45574, A
c	248	16	55.2	1866	8	US-10-750-623-45574 Sequence 45574, A
c	249	16	55.2	1997	6	US-09-925-065A-29187 Sequence 29187, A
c	250	16	55.2	1997	6	US-09-925-065A-29188 Sequence 29188, A
c	251	16	55.2	1997	9	US-10-301-480-130425 Sequence 130425,
c	252	16	55.2	1997	9	US-10-301-480-130426 Sequence 130426,
c	253	16	55.2	1997	10	US-10-301-480-743834 Sequence 743834,
c	254	16	55.2	1997	10	US-10-301-480-743835 Sequence 743835,
c	255	16	55.2	2098	6	US-09-925-065A-87472 Sequence 87472, A
c	256	16	55.2	2098	9	US-10-301-480-188713 Sequence 188713,
c	257	16	55.2	2098	10	US-10-301-480-802122 Sequence 802122,
	258	16	55.2	2147	6	US-09-925-065A-709005 Sequence 709005,
	259	16	55.2	2385	9	US-10-301-480-97257 Sequence 97257, A
	260	16	55.2	2385	9	US-10-301-480-97258 Sequence 97258, A
	261	16	55.2	2385	9	US-10-301-480-97259 Sequence 97259, A
	262	16	55.2	2385	9	US-10-301-480-97260 Sequence 97260, A
	263	16	55.2	2385	10	US-10-301-480-710666 Sequence 710666,
	264	16	55.2	2385	10	US-10-301-480-710667 Sequence 710667,
	265	16	55.2	2385	10	US-10-301-480-710668 Sequence 710668,
	266	16	55.2	2385	10	US-10-301-480-710669 Sequence 710669,
	267	16	55.2	2524	6	US-09-925-065A-672433 Sequence 672433,
	268	16	55.2	2524	6	US-09-925-065A-718111 Sequence 718111,
	269	16	55.2	3001	14	US-11-136-527-1015 Sequence 1015, Ap
	270	16	55.2	3187	11	US-11-072-512-1004 Sequence 1004, Ap
c	271	16	55.2	4891	7	US-10-204-639-136 Sequence 136, App
c	272	16	55.2	16082	8	US-10-995-561-13485 Sequence 13485, A
	273	16	55.2	46215	8	US-10-995-561-13483 Sequence 13483, A
c	274	16	55.2	46215	8	US-10-995-561-13483 Sequence 13483, A
c	275	16	55.2	100000	14	US-11-124-367A-5036 Sequence 5036, Ap
c	276	16	55.2	120096	14	US-11-121-086-24 Sequence 24, Appl
c	277	16	55.2	126552	14	US-11-121-086-1 Sequence 1, Appli
c	278	16	55.2	142303	14	US-11-121-086-42 Sequence 42, Appl
c	279	16	55.2	150468	14	US-11-112-908-56 Sequence 56, Appl
c	280	16	55.2	153752	9	US-10-330-773-508 Sequence 508, App
	281	16	55.2	165857	14	US-11-121-086-34 Sequence 34, Appl
	282	16	55.2	172111	14	US-11-121-086-28 Sequence 28, Appl
c	283	16	55.2	174009	7	US-10-960-414-490 Sequence 490, App
c	284	16	55.2	176503	14	US-11-121-086-53 Sequence 53, Appl
c	285	16	55.2	191684	14	US-11-121-086-2 Sequence 2, Appli
c	286	16	55.2	193789	14	US-11-112-908-55 Sequence 55, Appl
	287	16	55.2	198977	11	US-11-114-798-54 Sequence 54, Appl
c	288	16	55.2	215126	9	US-10-330-773-339 Sequence 339, App
	289	16	55.2	260209	8	US-10-933-025-23 Sequence 23, Appl
	290	16	55.2	260209	11	US-11-219-360-23 Sequence 23, Appl
c	291	16	55.2	1125000	8	US-10-995-561-13286 Sequence 13286, A
	292	15	51.7	23	8	US-10-310-914A-1280388 Sequence 1280388,
	293	15	51.7	24	8	US-10-310-914A-1280324 Sequence 1280324,
c	294	15	51.7	201	8	US-10-995-561-74322 Sequence 74322, A
c	295	15	51.7	201	14	US-11-124-368A-7903 Sequence 7903, Ap
c	296	15	51.7	301	6	US-09-925-065A-511728 Sequence 511728,
c	297	15	51.7	301	6	US-09-925-065A-511729 Sequence 511729,
c	298	15	51.7	345	6	US-09-925-065A-496756 Sequence 496756,
	299	15	51.7	371	6	US-09-925-065A-654890 Sequence 654890,
	300	15	51.7	371	10	US-10-301-480-347748 Sequence 347748,
	301	15	51.7	371	10	US-10-301-480-961157 Sequence 961157,
	302	15	51.7	377	6	US-09-925-065A-270241 Sequence 270241,
c	303	15	51.7	387	6	US-09-925-065A-238927 Sequence 238927,
c	304	15	51.7	402	6	US-09-925-065A-238926 Sequence 238926,
c	305	15	51.7	403	9	US-10-301-480-13322 Sequence 13322, A
c	306	15	51.7	403	10	US-10-301-480-626731 Sequence 626731,
	307	15	51.7	406	10	US-10-301-480-320953 Sequence 320953,
	308	15	51.7	406	10	US-10-301-480-934362 Sequence 934362,
c	309	15	51.7	407	6	US-09-925-065A-950450 Sequence 950450,
	310	15	51.7	427	10	US-10-301-480-320952 Sequence 320952,
	311	15	51.7	427	10	US-10-301-480-934361 Sequence 934361,
c	312	15	51.7	430	9	US-10-301-480-13323 Sequence 13323, A
c	313	15	51.7	430	10	US-10-301-480-626732 Sequence 626732,
	314	15	51.7	494	6	US-09-925-065A-639136 Sequence 639136,
	315	15	51.7	494	6	US-09-925-065A-639137 Sequence 639137,
	316	15	51.7	494	6	US-09-925-065A-639138 Sequence 639138,
	317	15	51.7	498	6	US-09-925-065A-624913 Sequence 624913,
	318	15	51.7	498	10	US-10-301-480-523837 Sequence 523837,
	319	15	51.7	498	10	US-10-301-480-523838 Sequence 523838,
	320	15	51.7	498	10	US-10-301-480-1137246 Sequence 1137246,
	321	15	51.7	498	10	US-10-301-480-1137247 Sequence 1137247,
c	322	15	51.7	499	6	US-09-925-065A-556777 Sequence 556777,
	323	15	51.7	514	6	US-09-925-065A-831590 Sequence 831590,
	324	15	51.7	525	6	US-09-925-065A-468385 Sequence 468385,
	325	15	51.7	525	6	US-09-925-065A-468386 Sequence 468386,
	326	15	51.7	526	6	US-09-925-065A-426860 Sequence 426860,
	327	15	51.7	535	6	US-09-925-065A-603411 Sequence 603411,
c	328	15	51.7	537	6	US-09-925-065A-92727 Sequence 92727, A
c	329	15	51.7	537	9	US-10-301-480-193969 Sequence 193969,
c	330	15	51.7	537	10	US-10-301-480-807378 Sequence 807378,
c	331	15	51.7	542	6	US-09-925-065A-356119 Sequence 356119,
c	332	15	51.7	542	6	US-09-925-065A-356120 Sequence 356120,
c	333	15	51.7	542	6	US-09-925-065A-356121 Sequence 356121,
	334	15	51.7	543	9	US-10-301-480-27518 Sequence 27518, A
	335	15	51.7	543	10	US-10-301-480-640927 Sequence 640927,
	336	15	51.7	546	6	US-09-925-065A-50642 Sequence 50642, A
	337	15	51.7	546	9	US-10-301-480-151880 Sequence 151880,
	338	15	51.7	546	10	US-10-301-480-765289 Sequence 765289,
	339	15	51.7	547	6	US-09-925-065A-212785 Sequence 212785,
c	340	15	51.7	551	10	US-10-301-480-427583 Sequence 427583,
c	341	15	51.7	551	10	US-10-301-480-427584 Sequence 427584,
c	342	15	51.7	551	10	US-10-301-480-427585 Sequence 427585,
c	343	15	51.7	551	10	US-10-301-480-1040992 Sequence 1040992,
c	344	15	51.7	551	10	US-10-301-480-1040993 Sequence 1040993,
c	345	15	51.7	551	10	US-10-301-480-1040994 Sequence 1040994,
c	346	15	51.7	553	6	US-09-925-065A-270337 Sequence 270337,
c	347	15	51.7	555	6	US-09-925-065A-481585 Sequence 481585,
c	348	15	51.7	561	10	US-10-301-480-299822 Sequence 299822,
c	349	15	51.7	561	10	US-10-301-480-913231 Sequence 913231,
c	350	15	51.7	564	6	US-09-925-065A-857894 Sequence 857894,
	351	15	51.7	566	6	US-09-925-065A-784618 Sequence 784618,
c	352	15	51.7	567	6	US-09-925-065A-46171 Sequence 46171, A
c	353	15	51.7	567	9	US-10-301-480-147409 Sequence 147409,
c	354	15	51.7	567	10	US-10-301-480-760818 Sequence 760818,
c	355	15	51.7	575	6	US-09-925-065A-270338 Sequence 270338,
c	356	15	51.7	575	6	US-09-925-065A-270339 Sequence 270339,
	357	15	51.7	577	6	US-09-925-065A-386623 Sequence 386623,
	358	15	51.7	580	6	US-09-925-065A-199580 Sequence 199580,
	359	15	51.7	581	6	US-09-925-065A-318965 Sequence 318965,
c	360	15	51.7	581	6	US-09-925-065A-319269 Sequence 319269,
c	361	15	51.7	585	6	US-09-925-065A-347150 Sequence 347150,
c	362	15	51.7	585	6	US-09-925-065A-347151 Sequence 347151,
	363	15	51.7	585	6	US-09-925-065A-955975 Sequence 955975,
	364	15	51.7	586	10	US-10-301-480-393761 Sequence 393761,
c	365	15	51.7	586	10	US-10-301-480-394033 Sequence 394033,
	366	15	51.7	586	10	US-10-301-480-1007170 Sequence 1007170,
c	367	15	51.7	586	10	US-10-301-480-1007442 Sequence 1007442,
	368	15	51.7	594	6	US-09-925-065A-680933 Sequence 680933,
	369	15	51.7	594	6	US-09-925-065A-799965 Sequence 799965,
c	370	15	51.7	596	6	US-09-925-065A-674125 Sequence 674125,
	371	15	51.7	596	10	US-10-301-480-288270 Sequence 288270,
	372	15	51.7	596	10	US-10-301-480-901679 Sequence 901679,
c	373	15	51.7	600	6	US-09-925-065A-652274 Sequence 652274,
c	374	15	51.7	604	6	US-09-925-065A-941993 Sequence 941993,
	375	15	51.7	605	6	US-09-925-065A-493602 Sequence 493602,
c	376	15	51.7	605	10	US-10-301-480-465729 Sequence 465729,
c	377	15	51.7	605	10	US-10-301-480-1079138 Sequence 1079138,
c	378	15	51.7	606	6	US-09-925-065A-348409 Sequence 348409,
	379	15	51.7	608	6	US-09-925-065A-398298 Sequence 398298,
	380	15	51.7	608	6	US-09-925-065A-935115 Sequence 935115,
c	381	15	51.7	611	6	US-09-925-065A-887092 Sequence 887092,
c	382	15	51.7	611	6	US-09-925-065A-887093 Sequence 887093,
	383	15	51.7	615	6	US-09-925-065A-21541 Sequence 21541, A

	384	15	51.7	615	9	US-10-301-480-122778 Sequence 122778,
	385	15	51.7	615	10	US-10-301-480-736187 Sequence 736187,
	386	15	51.7	617	6	US-09-925-065A-935317 Sequence 935317,
c	387	15	51.7	618	10	US-10-301-480-374743 Sequence 374743,
c	388	15	51.7	618	10	US-10-301-480-988152 Sequence 988152,
c	389	15	51.7	622	6	US-09-925-065A-298062 Sequence 298062,
	390	15	51.7	626	6	US-09-925-065A-920521 Sequence 920521,
	391	15	51.7	628	10	US-10-301-480-420552 Sequence 420552,
	392	15	51.7	628	10	US-10-301-480-1033961 Sequence 1033961,
c	393	15	51.7	632	10	US-10-301-480-378263 Sequence 378263,
c	394	15	51.7	632	10	US-10-301-480-378264 Sequence 378264,
c	395	15	51.7	632	10	US-10-301-480-991672 Sequence 991672,
c	396	15	51.7	632	10	US-10-301-480-991673 Sequence 991673,
	397	15	51.7	635	6	US-09-925-065A-301980 Sequence 301980,
	398	15	51.7	635	6	US-09-925-065A-301981 Sequence 301981,
	399	15	51.7	635	6	US-09-925-065A-941160 Sequence 941160,
c	400	15	51.7	638	6	US-09-925-065A-793409 Sequence 793409,
c	401	15	51.7	638	10	US-10-301-480-350405 Sequence 350405,
c	402	15	51.7	638	10	US-10-301-480-963814 Sequence 963814,
c	403	15	51.7	643	6	US-09-925-065A-273160 Sequence 273160,
	404	15	51.7	644	6	US-09-925-065A-940112 Sequence 940112,
c	405	15	51.7	648	10	US-10-301-480-230100 Sequence 230100,
c	406	15	51.7	648	10	US-10-301-480-843509 Sequence 843509,
	407	15	51.7	651	10	US-10-301-480-573584 Sequence 573584,
	408	15	51.7	651	10	US-10-301-480-1186993 Sequence 1186993,
c	409	15	51.7	660	6	US-09-925-065A-53162 Sequence 53162, A
c	410	15	51.7	660	9	US-10-301-480-154400 Sequence 154400,
c	411	15	51.7	660	10	US-10-301-480-767809 Sequence 767809,
c	412	15	51.7	662	6	US-09-925-065A-134359 Sequence 134359,
c	413	15	51.7	667	6	US-09-925-065A-921521 Sequence 921521,
c	414	15	51.7	668	6	US-09-925-065A-934679 Sequence 934679,
c	415	15	51.7	670	9	US-10-301-480-80750 Sequence 80750, A
c	416	15	51.7	670	9	US-10-301-480-80751 Sequence 80751, A
c	417	15	51.7	670	10	US-10-301-480-694159 Sequence 694159,
c	418	15	51.7	670	10	US-10-301-480-694160 Sequence 694160,
	419	15	51.7	690	10	US-10-301-480-605394 Sequence 605394,
	420	15	51.7	690	10	US-10-301-480-1218803 Sequence 1218803,
	421	15	51.7	699	10	US-10-301-480-607532 Sequence 607532,
	422	15	51.7	699	10	US-10-301-480-1220941 Sequence 1220941,
c	423	15	51.7	721	10	US-10-301-480-551610 Sequence 551610,
c	424	15	51.7	721	10	US-10-301-480-1165019 Sequence 1165019,
c	425	15	51.7	765	6	US-09-925-065A-6229 Sequence 6229, Ap
c	426	15	51.7	765	9	US-10-301-480-107466 Sequence 107466,
c	427	15	51.7	765	10	US-10-301-480-720875 Sequence 720875,
c	428	15	51.7	794	6	US-09-925-065A-701729 Sequence 701729,
c	429	15	51.7	900	8	US-10-750-185-42870 Sequence 42870, A
c	430	15	51.7	900	8	US-10-750-623-42870 Sequence 42870, A
	431	15	51.7	962	8	US-10-750-185-64851 Sequence 64851, A
	432	15	51.7	962	8	US-10-750-623-64851 Sequence 64851, A
c	433	15	51.7	964	10	US-10-301-480-601396 Sequence 601396,
c	434	15	51.7	964	10	US-10-301-480-1214805 Sequence 1214805,
c	435	15	51.7	991	10	US-10-301-480-555051 Sequence 555051,
c	436	15	51.7	991	10	US-10-301-480-1168460 Sequence 1168460,
c	437	15	51.7	994	10	US-10-301-480-611396 Sequence 611396,
c	438	15	51.7	994	10	US-10-301-480-1224805 Sequence 1224805,
	439	15	51.7	996	10	US-10-301-480-582690 Sequence 582690,
	440	15	51.7	996	10	US-10-301-480-1196099 Sequence 1196099,
	441	15	51.7	1061	6	US-09-925-065A-554557 Sequence 554557,
	442	15	51.7	1061	6	US-09-925-065A-554558 Sequence 554558,
	443	15	51.7	1228	8	US-10-750-185-56444 Sequence 56444, A
	444	15	51.7	1228	8	US-10-750-623-56444 Sequence 56444, A
c	445	15	51.7	1545	6	US-09-925-065A-689221 Sequence 689221,
c	446	15	51.7	1616	6	US-09-925-065A-62721 Sequence 62721, A
c	447	15	51.7	1616	9	US-10-301-480-163959 Sequence 163959,
c	448	15	51.7	1616	10	US-10-301-480-777368 Sequence 777368,
	449	15	51.7	1916	6	US-09-925-065A-54697 Sequence 54697, A
	450	15	51.7	1916	9	US-10-301-480-155935 Sequence 155935,
	451	15	51.7	1916	10	US-10-301-480-769344 Sequence 769344,
c	452	15	51.7	1947	6	US-09-925-065A-715529 Sequence 715529,
c	453	15	51.7	1947	6	US-09-925-065A-715530 Sequence 715530,
	454	15	51.7	2253	6	US-09-925-065A-85304 Sequence 85304, A
	455	15	51.7	2253	6	US-09-925-065A-85305 Sequence 85305, A
	456	15	51.7	2253	9	US-10-301-480-186544 Sequence 186544,
	457	15	51.7	2253	9	US-10-301-480-186545 Sequence 186545,
	458	15	51.7	2253	10	US-10-301-480-799953 Sequence 799953,
	459	15	51.7	2253	10	US-10-301-480-799954 Sequence 799954,
c	460	15	51.7	2415	6	US-09-925-065A-76261 Sequence 76261, A
c	461	15	51.7	2415	9	US-10-301-480-177500 Sequence 177500,
c	462	15	51.7	2415	10	US-10-301-480-790909 Sequence 790909,
c	463	15	51.7	3305	8	US-10-750-185-26892 Sequence 26892, A
c	464	15	51.7	3305	8	US-10-750-623-26892 Sequence 26892, A
c	465	15	51.7	12657	8	US-10-995-561-13454 Sequence 13454, A
c	466	15	51.7	41309	8	US-10-995-561-13487 Sequence 13487, A
c	467	15	51.7	45517	8	US-10-995-561-13455 Sequence 13455, A
	468	15	51.7	46854	14	US-11-124-368A-2892 Sequence 2892, Ap
	469	15	51.7	46878	8	US-10-995-561-13276 Sequence 13276, A
c	470	15	51.7	47285	14	US-11-124-367A-5096 Sequence 5096, Ap
	471	15	51.7	53332	8	US-10-786-065-3 Sequence 3, Appli
	472	15	51.7	76138	9	US-10-330-773-311 Sequence 311, App
c	473	15	51.7	95484	9	US-10-330-773-275 Sequence 275, App
	474	15	51.7	121160	9	US-10-330-773-847 Sequence 847, App
c	475	15	51.7	127340	14	US-11-112-908-35 Sequence 35, Appl
c	476	15	51.7	147700	8	US-10-857-780-3 Sequence 3, Appli
	477	15	51.7	153142	14	US-11-121-086-27 Sequence 27, Appl
	478	15	51.7	153376	14	US-11-121-086-5 Sequence 5, Appli
	479	15	51.7	154452	14	US-11-121-086-74 Sequence 74, Appl
c	480	15	51.7	165627	14	US-11-121-086-89 Sequence 89, Appl
	481	15	51.7	171486	14	US-11-121-086-105 Sequence 105, App
c	482	15	51.7	171486	14	US-11-121-086-105 Sequence 105, App
	483	15	51.7	172543	14	US-11-121-086-6 Sequence 6, Appli
c	484	15	51.7	173602	14	US-11-121-086-25 Sequence 25, Appl
c	485	15	51.7	178024	9	US-10-330-773-698 Sequence 698, App
	486	15	51.7	179666	14	US-11-121-086-67 Sequence 67, Appl
c	487	15	51.7	179777	14	US-11-121-086-106 Sequence 106, App
c	488	15	51.7	180862	14	US-11-112-908-40 Sequence 40, Appl
	489	15	51.7	181172	14	US-11-121-086-41 Sequence 41, Appl
c	490	15	51.7	193363	14	US-11-112-908-32 Sequence 32, Appl
	491	15	51.7	201309	11	US-11-114-798-51 Sequence 51, Appl
c	492	15	51.7	215248	9	US-10-330-773-261 Sequence 261, App
c	493	15	51.7	268685	8	US-10-933-025-22 Sequence 22, Appl
c	494	15	51.7	268685	11	US-11-219-360-22 Sequence 22, Appl
	495	15	51.7	272022	9	US-10-330-773-102 Sequence 102, App
	496	15	51.7	341511	9	US-10-472-808A-3 Sequence 3, Appli
	497	15	51.7	380749	8	US-10-995-561-13216 Sequence 13216, A
	498	15	51.7	1080000	8	US-10-928-446A-1 Sequence 1, Appli
c	499	15	51.7	1080000	8	US-10-928-446A-1 Sequence 1, Appli
	500	15	51.7	1080000	8	US-10-928-446A-181 Sequence 181, App

ALIGNMENTS

```
RESULT 1
US-09-925-065A-577180/c
; Sequence 577180, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 577180
```

```
;   LENGTH: 634
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-577180

  Query Match             58.6%;  Score 17;  DB 6;  Length 634;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db        434 TCAGGCATGAGCCAGCA 418


RESULT 2
US-09-925-065A-577181/c
; Sequence 577181, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 577181
;   LENGTH: 634
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-577181

  Query Match             58.6%;  Score 17;  DB 6;  Length 634;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db        434 TCAGGCATGAGCCAGCA 418


RESULT 3
US-09-925-065A-577182/c
; Sequence 577182, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 577182
;   LENGTH: 634
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-577182

  Query Match             58.6%;  Score 17;  DB 6;  Length 634;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 TCAGGCATGAGCCAGCA 29
              |||||||||||||||||
Db        434 TCAGGCATGAGCCAGCA 418


RESULT 4
US-10-330-773-362/c
; Sequence 362, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 362
;   LENGTH: 21728
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(21728)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-362

  Query Match             58.6%;  Score 17;  DB 9;  Length 21728;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAGG 17
              |||||||||||||||||
Db       5609 CCTCTCTGGAGCTCAGG 5593


RESULT 5
US-10-310-914A-1280357
; Sequence 1280357, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
;  APPLICANT: Bentwich, Isaac
;  APPLICANT:  Shiler, Kvuzat
;  TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes an
;  TITLE OF INVENTION:  uses thereof
;  FILE REFERENCE: 06087.0200.CPUS01
;  CURRENT APPLICATION NUMBER: US/10/310,914A
;  CURRENT FILING DATE:  2002-12-06
;  NUMBER OF SEQ ID NOS: 1388402
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1280357
;   LENGTH: 22
;   TYPE: RNA
;   ORGANISM: Human
US-10-310-914A-1280357

  Query Match             55.2%;  Score 16;  DB 8;  Length 22;
  Best Local Similarity   93.8%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          14 CAGGCATGAGCCAGCA 29
               ||||||:|||||||||
Db           6 CAGGCAUGAGCCAGCA 21


RESULT 6
US-10-310-914A-1280364
; Sequence 1280364, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
;  APPLICANT: Bentwich, Isaac
;  APPLICANT:  Shiler, Kvuzat
;  TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
;  TITLE OF INVENTION:  uses thereof
;  FILE REFERENCE: 06087.0200.CPUS01
;  CURRENT APPLICATION NUMBER: US/10/310,914A
;  CURRENT FILING DATE:  2002-12-06
;  NUMBER OF SEQ ID NOS: 1388402
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1280364
;   LENGTH: 22
;   TYPE: RNA
;   ORGANISM: Human
US-10-310-914A-1280364

  Query Match             55.2%;  Score 16;  DB 8;  Length 22;
  Best Local Similarity   93.8%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          14 CAGGCATGAGCCAGCA 29
               ||||||:|||||||||
Db           3 CAGGCAUGAGCCAGCA 18


RESULT 7
US-10-310-914A-1280360
; Sequence 1280360, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
;  APPLICANT: Bentwich, Isaac
;  APPLICANT:  Shiler, Kvuzat
;  TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
;  TITLE OF INVENTION:  uses thereof
;  FILE REFERENCE: 06087.0200.CPUS01
;  CURRENT APPLICATION NUMBER: US/10/310,914A
;  CURRENT FILING DATE:  2002-12-06
;  NUMBER OF SEQ ID NOS: 1388402
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1280360
;   LENGTH: 24
;   TYPE: RNA
;   ORGANISM: Human
US-10-310-914A-1280360

  Query Match             55.2%;  Score 16;  DB 8;  Length 24;
  Best Local Similarity   93.8%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          14 CAGGCATGAGCCAGCA 29
               ||||||:|||||||||
Db           7 CAGGCAUGAGCCAGCA 22


RESULT 8
US-10-310-914A-1280389
; Sequence 1280389, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
;  APPLICANT: Bentwich, Isaac
;  APPLICANT:  Shiler, Kvuzat
;  TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
;  TITLE OF INVENTION:  uses thereof
;  FILE REFERENCE: 06087.0200.CPUS01
;  CURRENT APPLICATION NUMBER: US/10/310,914A
;  CURRENT FILING DATE:  2002-12-06
;  NUMBER OF SEQ ID NOS: 1388402
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1280389
;   LENGTH: 24
;   TYPE: RNA
;   ORGANISM: Human
US-10-310-914A-1280389

  Query Match             55.2%;  Score 16;  DB 8;  Length 24;
  Best Local Similarity   93.8%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          14 CAGGCATGAGCCAGCA 29
               ||||||:|||||||||
Db           9 CAGGCAUGAGCCAGCA 24


RESULT 9
US-10-310-914A-1280390
; Sequence 1280390, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
;  APPLICANT: Bentwich, Isaac
;  APPLICANT:  Shiler, Kvuzat
;  TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
;  TITLE OF INVENTION:  uses thereof
;  FILE REFERENCE: 06087.0200.CPUS01
;  CURRENT APPLICATION NUMBER: US/10/310,914A
;  CURRENT FILING DATE:  2002-12-06
;  NUMBER OF SEQ ID NOS: 1388402
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1280390
;   LENGTH: 25
;   TYPE: RNA
;   ORGANISM: Human
US-10-310-914A-1280390

  Query Match             55.2%;  Score 16;  DB 8;  Length 25;
  Best Local Similarity   93.8%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          14 CAGGCATGAGCCAGCA 29
               ||||||:|||||||||
Db           9 CAGGCAUGAGCCAGCA 24


RESULT 10
US-11-136-527-241234
; Sequence 241234, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 241234
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: Probe
US-11-136-527-241234

  Query Match             55.2%;  Score 16;  DB 14;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+03;
```

```
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db          6 CAGGCATGAGCCAGCA 21


RESULT 11
US-11-136-527-241243
; Sequence 241243, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 241243
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: Probe
US-11-136-527-241243

  Query Match             55.2%;  Score 16;  DB 14;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db          1 CAGGCATGAGCCAGCA 16


RESULT 12
US-11-136-527-241246
; Sequence 241246, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 241246
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: Probe
US-11-136-527-241246

  Query Match             55.2%;  Score 16;  DB 14;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db          5 CAGGCATGAGCCAGCA 20


RESULT 13
US-10-310-914A-1280391
; Sequence 1280391, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
;  APPLICANT: Bentwich, Isaac
;  APPLICANT:  Shiler, Kvuzat
;  TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes ar
;  TITLE OF INVENTION:  uses thereof
;  FILE REFERENCE: 06087.0200.CPUS01
;  CURRENT APPLICATION NUMBER: US/10/310,914A
;  CURRENT FILING DATE:  2002-12-06
;  NUMBER OF SEQ ID NOS: 1388402
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1280391
;   LENGTH: 27
;   TYPE: RNA
;   ORGANISM: Human
US-10-310-914A-1280391

  Query Match             55.2%;  Score 16;  DB 8;  Length 27;
  Best Local Similarity   93.8%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||:|||||||||
Db          9 CAGGCAUGAGCCAGCA 24


RESULT 14
US-10-995-561-41746/c
; Sequence 41746, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41746
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-41746

  Query Match             55.2%;  Score 16;  DB 8;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 7.4e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db         73 CAGGCATGAGCCAGCA 58


RESULT 15
US-10-995-561-79446/c
; Sequence 79446, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 79446
;   LENGTH: 201
;   TYPE: DNA
```

```
;   ORGANISM: Homo sapiens
US-10-995-561-79446

  Query Match             55.2%;  Score 16;  DB 8;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 7.4e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        186 CAGGCATGAGCCAGCA 171


RESULT 16
US-10-995-561-79792/c
; Sequence 79792, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 79792
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-79792

  Query Match             55.2%;  Score 16;  DB 8;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 7.4e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        186 CAGGCATGAGCCAGCA 171


RESULT 17
US-11-124-367A-15249
; Sequence 15249, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  Hongjin Huang
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Fibrosis Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001519.ORD
;  CURRENT APPLICATION NUMBER: US/11/124,367A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,846
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/582,609
;  PRIOR FILING DATE: 2004-06-25
;  PRIOR APPLICATION NUMBER: US 60/599,554
;  PRIOR FILING DATE: 2004-08-09
;  NUMBER OF SEQ ID NOS: 34460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15249
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-124-367A-15249

  Query Match             55.2%;  Score 16;  DB 14;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 7.4e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTCTCTGGAGCTCAG 16
              ||||||||||||||||
Db        141 CCTCTCTGGAGCTCAG 156


RESULT 18
US-11-124-367A-29976/c
; Sequence 29976, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  Hongjin Huang
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Fibrosis Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001519.ORD
;  CURRENT APPLICATION NUMBER: US/11/124,367A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,846
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/582,609
;  PRIOR FILING DATE: 2004-06-25
;  PRIOR APPLICATION NUMBER: US 60/599,554
;  PRIOR FILING DATE: 2004-08-09
;  NUMBER OF SEQ ID NOS: 34460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29976
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-124-367A-29976

  Query Match             55.2%;  Score 16;  DB 14;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 7.4e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db         73 CAGGCATGAGCCAGCA 58


RESULT 19
US-10-301-480-207048
; Sequence 207048, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 207048
;   LENGTH: 293
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-207048

  Query Match             55.2%;  Score 16;  DB 9;  Length 293;
  Best Local Similarity   100.0%;  Pred. No. 6.9e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        112 CAGGCATGAGCCAGCA 127


RESULT 20
US-10-301-480-207050
; Sequence 207050, Application US/10301480
```

```
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 207050
;   LENGTH: 293
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-207050

  Query Match             55.2%;  Score 16;  DB 9;  Length 293;
  Best Local Similarity   100.0%;  Pred. No. 6.9e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        112 CAGGCATGAGCCAGCA 127


RESULT 21
US-10-301-480-820457
; Sequence 820457, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 820457
;   LENGTH: 293
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-820457

  Query Match             55.2%;  Score 16;  DB 10;  Length 293;
  Best Local Similarity   100.0%;  Pred. No. 6.9e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        112 CAGGCATGAGCCAGCA 127


RESULT 22
US-10-301-480-820459
; Sequence 820459, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 820459
;   LENGTH: 293
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-820459

  Query Match             55.2%;  Score 16;  DB 10;  Length 293;
  Best Local Similarity   100.0%;  Pred. No. 6.9e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        112 CAGGCATGAGCCAGCA 127


RESULT 23
US-09-925-065A-107140
; Sequence 107140, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 107140
;   LENGTH: 299
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-107140

  Query Match             55.2%;  Score 16;  DB 6;  Length 299;
  Best Local Similarity   100.0%;  Pred. No. 6.8e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        112 CAGGCATGAGCCAGCA 127


RESULT 24
US-09-925-065A-107142
; Sequence 107142, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
```

```
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 107142
;   LENGTH: 299
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-107142

  Query Match             55.2%;  Score 16;  DB 6;  Length 299;
  Best Local Similarity   100.0%;  Pred. No. 6.8e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        112 CAGGCATGAGCCAGCA 127


RESULT 25
US-09-925-065A-431188/c
; Sequence 431188, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 431188
;   LENGTH: 370
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-431188

  Query Match             55.2%;  Score 16;  DB 6;  Length 370;
  Best Local Similarity   100.0%;  Pred. No. 6.6e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db         71 CAGGCATGAGCCAGCA 56


RESULT 26
US-10-301-480-493877
; Sequence 493877, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 493877
;   LENGTH: 373
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-493877

  Query Match             55.2%;  Score 16;  DB 10;  Length 373;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        302 CAGGCATGAGCCAGCA 317


RESULT 27
US-10-301-480-1107286
; Sequence 1107286, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1107286
;   LENGTH: 373
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1107286

  Query Match             55.2%;  Score 16;  DB 10;  Length 373;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
               ||||||||||||||||
Db        302 CAGGCATGAGCCAGCA 317


RESULT 28
US-10-301-480-252290/c
; Sequence 252290, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 252290
;   LENGTH: 383
;   TYPE: DNA
;   ORGANISM: Homo sapien
```

```
US-10-301-480-252290

  Query Match             55.2%;  Score 16;  DB 10;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        164 CAGGCATGAGCCAGCA 149


RESULT 29
US-10-301-480-865699/c
; Sequence 865699, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 865699
;   LENGTH: 383
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-865699

  Query Match             55.2%;  Score 16;  DB 10;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        164 CAGGCATGAGCCAGCA 149


RESULT 30
US-10-301-480-252286/c
; Sequence 252286, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 252286
;   LENGTH: 384
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-252286

  Query Match             55.2%;  Score 16;  DB 10;  Length 384;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        165 CAGGCATGAGCCAGCA 150


RESULT 31
US-10-301-480-865695/c
; Sequence 865695, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 865695
;   LENGTH: 384
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-865695

  Query Match             55.2%;  Score 16;  DB 10;  Length 384;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        165 CAGGCATGAGCCAGCA 150


RESULT 32
US-09-925-065A-159403/c
; Sequence 159403, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159403
;   LENGTH: 386
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159403

  Query Match             55.2%;  Score 16;  DB 6;  Length 386;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        164 CAGGCATGAGCCAGCA 149


RESULT 33
```

```
US-09-925-065A-159399/c
; Sequence 159399, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159399
;   LENGTH: 387
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159399

  Query Match             55.2%;  Score 16;  DB 6;  Length 387;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        165 CAGGCATGAGCCAGCA 150


RESULT 34
US-10-301-480-204513/c
; Sequence 204513, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 204513
;   LENGTH: 388
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-204513

  Query Match             55.2%;  Score 16;  DB 9;  Length 388;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        169 CAGGCATGAGCCAGCA 154


RESULT 35
US-10-301-480-252293/c
; Sequence 252293, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 252293
;   LENGTH: 388
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-252293

  Query Match             55.2%;  Score 16;  DB 10;  Length 388;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        169 CAGGCATGAGCCAGCA 154


RESULT 36
US-10-301-480-817922/c
; Sequence 817922, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 817922
;   LENGTH: 388
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-817922

  Query Match             55.2%;  Score 16;  DB 10;  Length 388;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        169 CAGGCATGAGCCAGCA 154


RESULT 37
US-10-301-480-865702/c
; Sequence 865702, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
```

```
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 865702
;   LENGTH: 388
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-865702

  Query Match             55.2%;  Score 16;  DB 10;  Length 388;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        169 CAGGCATGAGCCAGCA 154


RESULT 38
US-09-925-065A-104346/c
; Sequence 104346, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 104346
;   LENGTH: 391
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-104346

  Query Match             55.2%;  Score 16;  DB 6;  Length 391;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        169 CAGGCATGAGCCAGCA 154


RESULT 39
US-09-925-065A-159406/c
; Sequence 159406, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159406
;   LENGTH: 391
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159406

  Query Match             55.2%;  Score 16;  DB 6;  Length 391;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        169 CAGGCATGAGCCAGCA 154


RESULT 40
US-10-301-480-252287/c
; Sequence 252287, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 252287
;   LENGTH: 392
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-252287

  Query Match             55.2%;  Score 16;  DB 10;  Length 392;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        173 CAGGCATGAGCCAGCA 158


RESULT 41
US-10-301-480-865696/c
; Sequence 865696, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 865696
;   LENGTH: 392
```

```
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-865696

  Query Match             55.2%;  Score 16;  DB 10;  Length 392;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        173 CAGGCATGAGCCAGCA 158


RESULT 42
US-09-925-065A-159400/c
; Sequence 159400, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159400
;   LENGTH: 395
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-159400

  Query Match             55.2%;  Score 16;  DB 6;  Length 395;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        173 CAGGCATGAGCCAGCA 158


RESULT 43
US-09-925-065A-173155/c
; Sequence 173155, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 173155
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-173155

  Query Match             55.2%;  Score 16;  DB 6;  Length 396;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        173 CAGGCATGAGCCAGCA 158


RESULT 44
US-10-301-480-264760/c
; Sequence 264760, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 264760
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-264760

  Query Match             55.2%;  Score 16;  DB 10;  Length 396;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        173 CAGGCATGAGCCAGCA 158


RESULT 45
US-10-301-480-878169/c
; Sequence 878169, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 878169
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-878169

  Query Match             55.2%;  Score 16;  DB 10;  Length 396;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
```

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        173 CAGGCATGAGCCAGCA 158
```

RESULT 46
US-09-925-065A-173157/c
; Sequence 173157, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 173157
; LENGTH: 400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-173157

Query Match 55.2%; Score 16; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        175 CAGGCATGAGCCAGCA 160
```

RESULT 47
US-09-925-065A-173158/c
; Sequence 173158, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 173158
; LENGTH: 400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-173158

Query Match 55.2%; Score 16; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        175 CAGGCATGAGCCAGCA 160
```

RESULT 48
US-09-925-065A-173159/c
; Sequence 173159, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 173159
; LENGTH: 400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-173159

Query Match 55.2%; Score 16; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        175 CAGGCATGAGCCAGCA 160
```

RESULT 49
US-10-301-480-204515/c
; Sequence 204515, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 204515
; LENGTH: 400
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-204515

Query Match 55.2%; Score 16; DB 9; Length 400;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;

```
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        143 CAGGCATGAGCCAGCA 128


RESULT 50
US-10-301-480-264762/c
; Sequence 264762, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 264762
;   LENGTH: 400
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-264762

  Query Match             55.2%;  Score 16;  DB 10;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CAGGCATGAGCCAGCA 29
              ||||||||||||||||
Db        175 CAGGCATGAGCCAGCA 160


Search completed: April 12, 2006, 14:21:48
Job time : 690 secs
```